

9 | HYPOTHESIS TESTING WITH ONE SAMPLE



Figure 9.1 You can use a hypothesis test to decide if a dog breeder's claim that every Dalmatian has 35 spots is statistically sound. (Credit: Robert Neff)

Introduction

Chapter Objectives

By the end of this chapter, the student should be able to:

- Differentiate between Type I and Type II Errors
- Describe hypothesis testing in general and in practice
- Conduct and interpret hypothesis tests for a single population mean, population standard deviation known.
- Conduct and interpret hypothesis tests for a single population mean, population standard deviation unknown.
- Conduct and interpret hypothesis tests for a single population proportion.

One job of a statistician is to make statistical inferences about populations based on samples taken from the population. **Confidence intervals** are one way to estimate a population parameter. Another way to make a statistical inference is to

make a decision about a parameter. For instance, a car dealer advertises that its new small truck gets 35 miles per gallon, on average. A tutoring service claims that its method of tutoring helps 90% of its students get an A or a B. A company says that women managers in their company earn an average of \$60,000 per year.

A statistician will make a decision about these claims. This process is called "**hypothesis testing**." A hypothesis test involves collecting data from a sample and evaluating the data. Then, the statistician makes a decision as to whether or not there is sufficient evidence, based upon analyses of the data, to reject the null hypothesis.

In this chapter, you will conduct hypothesis tests on single means and single proportions. You will also learn about the errors associated with these tests.

Hypothesis testing consists of two contradictory hypotheses or statements, a decision based on the data, and a conclusion. To perform a hypothesis test, a statistician will:

1. Set up two contradictory hypotheses.
2. Collect sample data (in homework problems, the data or summary statistics will be given to you).
3. Determine the correct distribution to perform the hypothesis test.
4. Analyze sample data by performing the calculations that ultimately will allow you to reject or decline to reject the null hypothesis.
5. Make a decision and write a meaningful conclusion.

NOTE

To do the hypothesis test homework problems for this chapter and later chapters, make copies of the appropriate special solution sheets. See [Appendix E](#).

9.1 | Null and Alternative Hypotheses

The actual test begins by considering two **hypotheses**. They are called the **null hypothesis** and the **alternative hypothesis**. These hypotheses contain opposing viewpoints.

H_0 : The null hypothesis: It is a statement about the population that either is believed to be true or is used to put forth an argument unless it can be shown to be incorrect beyond a reasonable doubt.

H_a : The alternative hypothesis: It is a claim about the population that is contradictory to H_0 and what we conclude when we reject H_0 .

Since the null and alternative hypotheses are contradictory, you must examine evidence to decide if you have enough evidence to reject the null hypothesis or not. The evidence is in the form of sample data.

After you have determined which hypothesis the sample supports, you make a **decision**. There are two options for a decision. They are "reject H_0 " if the sample information favors the alternative hypothesis or "do not reject H_0 " or "decline to reject H_0 " if the sample information is insufficient to reject the null hypothesis.

Mathematical Symbols Used in H_0 and H_a :

H_0	H_a
equal (=)	not equal (\neq) or greater than ($>$) or less than ($<$)
greater than or equal to (\geq)	less than ($<$)
less than or equal to (\leq)	more than ($>$)

Table 9.1

NOTE

H_0 always has a symbol with an equal in it. H_a never has a symbol with an equal in it. The choice of symbol depends on the wording of the hypothesis test. However, be aware that many researchers (including one of the co-authors in research work) use = in the null hypothesis, even with $>$ or $<$ as the symbol in the alternative hypothesis. This practice is acceptable because we only make the decision to reject or not reject the null hypothesis.

Example 9.1

H_0 : No more than 30% of the registered voters in Santa Clara County voted in the primary election. $p \leq 30$

H_a : More than 30% of the registered voters in Santa Clara County voted in the primary election. $p > 30$

Try It

9.1 A medical trial is conducted to test whether or not a new medicine reduces cholesterol by 25%. State the null and alternative hypotheses.

Example 9.2

We want to test whether the mean GPA of students in American colleges is different from 2.0 (out of 4.0). The null and alternative hypotheses are:

$H_0: \mu = 2.0$

$H_a: \mu \neq 2.0$

Try It

9.2 We want to test whether the mean height of eighth graders is 66 inches. State the null and alternative hypotheses. Fill in the correct symbol ($=$, \neq , \geq , $<$, \leq , $>$) for the null and alternative hypotheses.

a. $H_0: \mu \underline{\hspace{0.5cm}} 66$

b. $H_a: \mu \underline{\hspace{0.5cm}} 66$

Example 9.3

We want to test if college students take less than five years to graduate from college, on the average. The null and alternative hypotheses are:

$H_0: \mu \geq 5$

$H_a: \mu < 5$

Try It

9.3 We want to test if it takes fewer than 45 minutes to teach a lesson plan. State the null and alternative hypotheses. Fill in the correct symbol ($=$, \neq , \geq , $<$, \leq , $>$) for the null and alternative hypotheses.

a. $H_0: \mu \underline{\hspace{0.5cm}} 45$

b. $H_a: \mu \underline{\hspace{0.5cm}} 45$

Example 9.4

In an issue of *U. S. News and World Report*, an article on school standards stated that about half of all students in France, Germany, and Israel take advanced placement exams and a third pass. The same article stated that 6.6% of U.S. students take advanced placement exams and 4.4% pass. Test if the percentage of U.S. students who take advanced placement exams is more than 6.6%. State the null and alternative hypotheses.

$$H_0: p \leq 0.066$$

$$H_a: p > 0.066$$

Try It

9.4 On a state driver's test, about 40% pass the test on the first try. We want to test if more than 40% pass on the first try. Fill in the correct symbol ($=$, \neq , \geq , $<$, \leq , $>$) for the null and alternative hypotheses.

a. $H_0: p \underline{\hspace{0.5cm}} 0.40$

b. $H_a: p \underline{\hspace{0.5cm}} 0.40$

Collaborative Exercise

Bring to class a newspaper, some news magazines, and some Internet articles. In groups, find articles from which your group can write null and alternative hypotheses. Discuss your hypotheses with the rest of the class.

9.2 | Outcomes and the Type I and Type II Errors

When you perform a hypothesis test, there are four possible outcomes depending on the actual truth (or falseness) of the null hypothesis H_0 and the decision to reject or not. The outcomes are summarized in the following table:

ACTION	H_0 IS ACTUALLY	...
	True	False
Do not reject H_0	Correct Outcome	Type II error
Reject H_0	Type I Error	Correct Outcome

Table 9.2

The four possible outcomes in the table are:

1. The decision is **not to reject H_0** when **H_0 is true (correct decision)**.
2. The decision is to **reject H_0** when **H_0 is true** (incorrect decision known as a **Type I error**).
3. The decision is **not to reject H_0** when, in fact, **H_0 is false** (incorrect decision known as a **Type II error**).
4. The decision is to **reject H_0** when **H_0 is false (correct decision** whose probability is called the **Power of the Test**).

Each of the errors occurs with a particular probability. The Greek letters α and β represent the probabilities.

α = probability of a Type I error = **$P(\text{Type I error})$** = probability of rejecting the null hypothesis when the null hypothesis is true.

β = probability of a Type II error = **$P(\text{Type II error})$** = probability of not rejecting the null hypothesis when the null hypothesis is false.

α and β should be as small as possible because they are probabilities of errors. They are rarely zero.

The Power of the Test is $1 - \beta$. Ideally, we want a high power that is as close to one as possible. Increasing the sample size can increase the Power of the Test.

The following are examples of Type I and Type II errors.

Example 9.5

Suppose the null hypothesis, H_0 , is: Frank's rock climbing equipment is safe.

Type I error: Frank thinks that his rock climbing equipment may not be safe when, in fact, it really is safe. **Type**

II error: Frank thinks that his rock climbing equipment may be safe when, in fact, it is not safe.

α = **probability** that Frank thinks his rock climbing equipment may not be safe when, in fact, it really is safe. β = **probability** that Frank thinks his rock climbing equipment may be safe when, in fact, it is not safe.

Notice that, in this case, the error with the greater consequence is the Type II error. (If Frank thinks his rock climbing equipment is safe, he will go ahead and use it.)

Try It

9.5 Suppose the null hypothesis, H_0 , is: the blood cultures contain no traces of pathogen X. State the Type I and Type II errors.

Example 9.6

Suppose the null hypothesis, H_0 , is: The victim of an automobile accident is alive when he arrives at the emergency room of a hospital.

Type I error: The emergency crew thinks that the victim is dead when, in fact, the victim is alive. **Type II error:** The emergency crew does not know if the victim is alive when, in fact, the victim is dead.

α = **probability** that the emergency crew thinks the victim is dead when, in fact, he is really alive = $P(\text{Type I error})$. β = **probability** that the emergency crew does not know if the victim is alive when, in fact, the victim is dead = $P(\text{Type II error})$.

The error with the greater consequence is the Type I error. (If the emergency crew thinks the victim is dead, they will not treat him.)

Try It

9.6 Suppose the null hypothesis, H_0 , is: a patient is not sick. Which type of error has the greater consequence, Type I or Type II?

Example 9.7

It's a Boy Genetic Labs claim to be able to increase the likelihood that a pregnancy will result in a boy being born. Statisticians want to test the claim. Suppose that the null hypothesis, H_0 , is: It's a Boy Genetic Labs has no effect on gender outcome.

Type I error: This results when a true null hypothesis is rejected. In the context of this scenario, we would state that we believe that It's a Boy Genetic Labs influences the gender outcome, when in fact it has no effect. The probability of this error occurring is denoted by the Greek letter alpha, α .

Type II error: This results when we fail to reject a false null hypothesis. In context, we would state that It's a Boy Genetic Labs does not influence the gender outcome of a pregnancy when, in fact, it does. The probability of this error occurring is denoted by the Greek letter beta, β .

The error of greater consequence would be the Type I error since couples would use the It's a Boy Genetic Labs product in hopes of increasing the chances of having a boy.

Try It

9.7 “Red tide” is a bloom of poison-producing algae—a few different species of a class of plankton called dinoflagellates. When the weather and water conditions cause these blooms, shellfish such as clams living in the area develop dangerous levels of a paralysis-inducing toxin. In Massachusetts, the Division of Marine Fisheries (DMF) monitors levels of the toxin in shellfish by regular sampling of shellfish along the coastline. If the mean level of toxin in clams exceeds 800 µg (micrograms) of toxin per kg of clam meat in any area, clam harvesting is banned there until the bloom is over and levels of toxin in clams subside. Describe both a Type I and a Type II error in this context, and state which error has the greater consequence.

Example 9.8

A certain experimental drug claims a cure rate of at least 75% for males with prostate cancer. Describe both the Type I and Type II errors in context. Which error is the more serious?

Type I: A cancer patient believes the cure rate for the drug is less than 75% when it actually is at least 75%.

Type II: A cancer patient believes the experimental drug has at least a 75% cure rate when it has a cure rate that is less than 75%.

In this scenario, the Type II error contains the more severe consequence. If a patient believes the drug works at least 75% of the time, this most likely will influence the patient’s (and doctor’s) choice about whether to use the drug as a treatment option.

Try It

9.8 Determine both Type I and Type II errors for the following scenario:

Assume a null hypothesis, H_0 , that states the percentage of adults with jobs is at least 88%.

Identify the Type I and Type II errors from these four statements.

- Not to reject the null hypothesis that the percentage of adults who have jobs is at least 88% when that percentage is actually less than 88%
- Not to reject the null hypothesis that the percentage of adults who have jobs is at least 88% when the percentage is actually at least 88%.
- Reject the null hypothesis that the percentage of adults who have jobs is at least 88% when the percentage is actually at least 88%.
- Reject the null hypothesis that the percentage of adults who have jobs is at least 88% when that percentage is actually less than 88%.

9.3 | Distribution Needed for Hypothesis Testing

Earlier in the course, we discussed sampling distributions. **Particular distributions are associated with hypothesis testing.** Perform tests of a population mean using a **normal distribution** or a **Student’s t -distribution**. (Remember, use a Student’s t -distribution when the population **standard deviation** is unknown and the distribution of the sample mean is approximately normal.) We perform tests of a population proportion using a normal distribution (usually n is large or the sample size is large).

If you are testing a **single population mean**, the distribution for the test is for **means**:

$$\bar{X} \sim N\left(\mu_X, \frac{\sigma_X}{\sqrt{n}}\right) \text{ or } t_{df}$$

The population parameter is μ . The estimated value (point estimate) for μ is \bar{x} , the sample mean.

If you are testing a **single population proportion**, the distribution for the test is for proportions or percentages:

$$P' \sim N\left(p, \sqrt{\frac{p \cdot q}{n}}\right)$$

The population parameter is p . The estimated value (point estimate) for p is p' . $p' = \frac{x}{n}$ where x is the number of successes and n is the sample size.

Assumptions

When you perform a **hypothesis test of a single population mean μ** using a **Student's t -distribution** (often called a t -test), there are fundamental assumptions that need to be met in order for the test to work properly. Your data should be a **simple random sample** that comes from a population that is approximately **normally distributed**. You use the sample **standard deviation** to approximate the population standard deviation. (Note that if the sample size is sufficiently large, a t -test will work even if the population is not approximately normally distributed).

When you perform a **hypothesis test of a single population mean μ** using a normal distribution (often called a z -test), you take a simple random sample from the population. The population you are testing is normally distributed or your sample size is sufficiently large. You know the value of the population standard deviation which, in reality, is rarely known.

When you perform a **hypothesis test of a single population proportion p** , you take a simple random sample from the population. You must meet the conditions for a **binomial distribution** which are: there are a certain number n of independent trials, the outcomes of any trial are success or failure, and each trial has the same probability of a success p . The shape of the binomial distribution needs to be similar to the shape of the normal distribution. To ensure this, the quantities np and nq must both be greater than five ($np > 5$ and $nq > 5$). Then the binomial distribution of a sample (estimated) proportion can be approximated by the normal distribution with $\mu = p$ and $\sigma = \sqrt{\frac{pq}{n}}$. Remember that $q = 1 - p$.

9.4 | Rare Events, the Sample, Decision and Conclusion

Establishing the type of distribution, sample size, and known or unknown standard deviation can help you figure out how to go about a hypothesis test. However, there are several other factors you should consider when working out a hypothesis test.

Rare Events

Suppose you make an assumption about a property of the population (this assumption is the **null hypothesis**). Then you gather sample data randomly. If the sample has properties that would be very **unlikely** to occur if the assumption is true, then you would conclude that your assumption about the population is probably incorrect. (Remember that your assumption is just an **assumption**—it is not a fact and it may or may not be true. But your sample data are real and the data are showing you a fact that seems to contradict your assumption.)

For example, Didi and Ali are at a birthday party of a very wealthy friend. They hurry to be first in line to grab a prize from a tall basket that they cannot see inside because they will be blindfolded. There are 200 plastic bubbles in the basket and Didi and Ali have been told that there is only one with a \$100 bill. Didi is the first person to reach into the basket and pull out a bubble. Her bubble contains a \$100 bill. The probability of this happening is $\frac{1}{200} = 0.005$. Because this is so unlikely,

Ali is hoping that what the two of them were told is wrong and there are more \$100 bills in the basket. A "rare event" has occurred (Didi getting the \$100 bill) so Ali doubts the assumption about only one \$100 bill being in the basket.

Using the Sample to Test the Null Hypothesis

Use the sample data to calculate the actual probability of getting the test result, called the **p -value**. The p -value is the **probability that, if the null hypothesis is true, the results from another randomly selected sample will be as extreme or more extreme as the results obtained from the given sample**.

A large p -value calculated from the data indicates that we should not reject the **null hypothesis**. The smaller the p -value, the more unlikely the outcome, and the stronger the evidence is against the null hypothesis. We would reject the null hypothesis if the evidence is strongly against it.

Draw a graph that shows the p -value. The hypothesis test is easier to perform if you use a graph because you see the problem more clearly.

Example 9.9

Suppose a baker claims that his bread height is more than 15 cm, on average. Several of his customers do not believe him. To persuade his customers that he is right, the baker decides to do a hypothesis test. He bakes 10 loaves of bread. The mean height of the sample loaves is 17 cm. The baker knows from baking hundreds of loaves of bread that the **standard deviation** for the height is 0.5 cm. and the distribution of heights is normal.

The null hypothesis could be $H_0: \mu \leq 15$ The alternate hypothesis is $H_a: \mu > 15$

The words "**is more than**" translates as a ">" so " $\mu > 15$ " goes into the alternate hypothesis. The null hypothesis must contradict the alternate hypothesis.

Since σ is known ($\sigma = 0.5$ cm.), the distribution for the population is known to be normal with mean $\mu = 15$ and standard deviation $\frac{\sigma}{\sqrt{n}} = \frac{0.5}{\sqrt{10}} = 0.16$.

Suppose the null hypothesis is true (the mean height of the loaves is no more than 15 cm). Then is the mean height (17 cm) calculated from the sample unexpectedly large? The hypothesis test works by asking the question how **unlikely** the sample mean would be if the null hypothesis were true. The graph shows how far out the sample mean is on the normal curve. The p -value is the probability that, if we were to take other samples, any other sample mean would fall at least as far out as 17 cm.

The p -value, then, is the probability that a sample mean is the same or greater than 17 cm. when the population mean is, in fact, 15 cm. We can calculate this probability using the normal distribution for means.

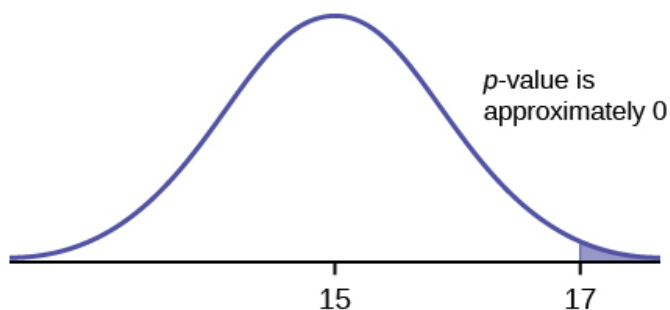


Figure 9.2

$p\text{-value} = P(\bar{x} > 17)$ which is approximately zero.

A p -value of approximately zero tells us that it is highly unlikely that a loaf of bread rises no more than 15 cm, on average. That is, almost 0% of all loaves of bread would be at least as high as 17 cm. **purely by CHANCE** had the population mean height really been 15 cm. Because the outcome of 17 cm. is so **unlikely (meaning it is happening NOT by chance alone)**, we conclude that the evidence is strongly against the null hypothesis (the mean height is at most 15 cm.). There is sufficient evidence that the true mean height for the population of the baker's loaves of bread is greater than 15 cm.

Try It Σ

9.9 A normal distribution has a standard deviation of 1. We want to verify a claim that the mean is greater than 12. A sample of 36 is taken with a sample mean of 12.5.

$H_0: \mu \leq 12$

$H_a: \mu > 12$

The p -value is 0.0013

Draw a graph that shows the p -value.

Decision and Conclusion

A systematic way to make a decision of whether to reject or not reject the **null hypothesis** is to compare the p -value and a **preset or preconceived α (also called a "significance level")**. A preset α is the probability of a **Type I error** (rejecting the null hypothesis when the null hypothesis is true). It may or may not be given to you at the beginning of the problem.

When you make a **decision** to reject or not reject H_0 , do as follows:

- If $\alpha > p$ -value, reject H_0 . The results of the sample data are significant. There is sufficient evidence to conclude that H_0 is an incorrect belief and that the **alternative hypothesis**, H_a , may be correct.
- If $\alpha \leq p$ -value, do not reject H_0 . The results of the sample data are not significant. There is not sufficient evidence to conclude that the alternative hypothesis, H_a , may be correct.
- When you "do not reject H_0 ", it does not mean that you should believe that H_0 is true. It simply means that the sample data have **failed** to provide sufficient evidence to cast serious doubt about the truthfulness of H_0 .

Conclusion: After you make your decision, write a thoughtful **conclusion** about the hypotheses in terms of the given problem.

Example 9.10

When using the p -value to evaluate a hypothesis test, it is sometimes useful to use the following memory device

If the p -value is low, the null must go.

If the p -value is high, the null must fly.

This memory aid relates a p -value less than the established alpha (the p is low) as rejecting the null hypothesis and, likewise, relates a p -value higher than the established alpha (the p is high) as not rejecting the null hypothesis.

Fill in the blanks.

Reject the null hypothesis when _____.

The results of the sample data _____.

Do not reject the null hypothesis when _____.

The results of the sample data _____.

Solution 9.10

Reject the null hypothesis when **the p -value is less than the established alpha value**. The results of the sample data **support the alternative hypothesis**.

Do not reject the null hypothesis when **the p -value is greater than the established alpha value**. The results of the sample data **do not support the alternative hypothesis**.

Try It

9.10 It's a Boy Genetics Labs claim their procedures improve the chances of a boy being born. The results for a test of a single population proportion are as follows:

$$H_0: p = 0.50, H_a: p > 0.50$$

$$\alpha = 0.01$$

$$p\text{-value} = 0.025$$

Interpret the results and state a conclusion in simple, non-technical terms.

9.5 | Additional Information and Full Hypothesis Test Examples

- In a **hypothesis test** problem, you may see words such as "the level of significance is 1%." The "1%" is the preconceived or preset α .
- The statistician setting up the hypothesis test selects the value of α to use **before** collecting the sample data.
- **If no level of significance is given, a common standard to use is $\alpha = 0.05$.**
- When you calculate the p -value and draw the picture, the p -value is the area in the left tail, the right tail, or split evenly between the two tails. For this reason, we call the hypothesis test left, right, or two tailed.
- The **alternative hypothesis**, H_a , tells you if the test is left, right, or two-tailed. It is the **key** to conducting the appropriate test.
- H_a **never** has a symbol that contains an equal sign.
- **Thinking about the meaning of the p -value:** A data analyst (and anyone else) should have more confidence that he made the correct decision to reject the null hypothesis with a smaller p -value (for example, 0.001 as opposed to 0.04) even if using the 0.05 level for alpha. Similarly, for a large p -value such as 0.4, as opposed to a p -value of 0.056 (alpha = 0.05 is less than either number), a data analyst should have more confidence that she made the correct decision in not rejecting the null hypothesis. This makes the data analyst use judgment rather than mindlessly applying rules.

The following examples illustrate a left-, right-, and two-tailed test.

Example 9.11

$$H_0: \mu = 5, H_a: \mu < 5$$

Test of a single population mean. H_a tells you the test is left-tailed. The picture of the p -value is as follows:

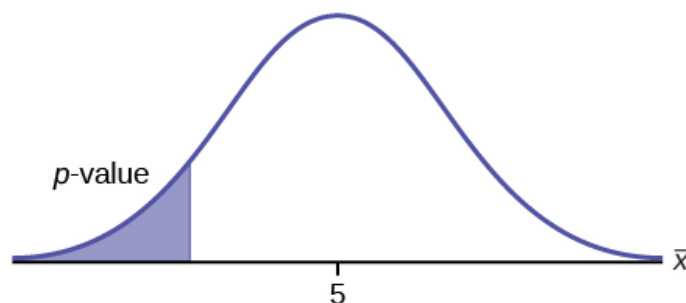


Figure 9.3

Try It Σ

9.11 $H_0: \mu = 10, H_a: \mu < 10$

Assume the p -value is 0.0935. What type of test is this? Draw the picture of the p -value.

Example 9.12

$$H_0: p \leq 0.2 \quad H_a: p > 0.2$$

This is a test of a single population proportion. H_a tells you the test is **right-tailed**. The picture of the p -value is as follows:

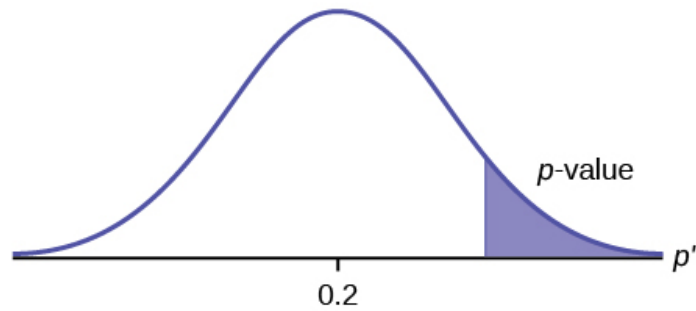


Figure 9.4

Try It Σ

9.12 $H_0: \mu \leq 1, H_a: \mu > 1$

Assume the $p\text{-value}$ is 0.1243. What type of test is this? Draw the picture of the $p\text{-value}$.

Example 9.13

$H_0: p = 50 \quad H_a: p \neq 50$

This is a test of a single population mean. H_a tells you the test is **two-tailed**. The picture of the $p\text{-value}$ is as follows.

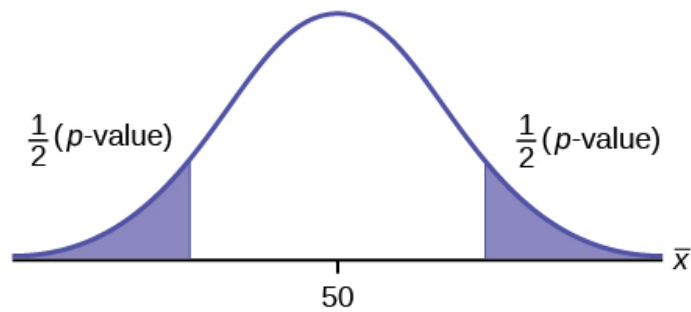


Figure 9.5

Try It Σ

9.13 $H_0: p = 0.5, H_a: p \neq 0.5$

Assume the $p\text{-value}$ is 0.2564. What type of test is this? Draw the picture of the $p\text{-value}$.

Full Hypothesis Test Examples

Example 9.14

Jeffrey, as an eight-year old, **established a mean time of 16.43 seconds** for swimming the 25-yard freestyle, with a **standard deviation of 0.8 seconds**. His dad, Frank, thought that Jeffrey could swim the 25-yard freestyle faster using goggles. Frank bought Jeffrey a new pair of expensive goggles and timed Jeffrey for **15 25-yard freestyle swims**. For the 15 swims, **Jeffrey's mean time was 16 seconds**. **Frank thought that the goggles helped Jeffrey to swim faster than the 16.43 seconds**. Conduct a hypothesis test using a preset $\alpha = 0.05$. Assume that the swim times for the 25-yard freestyle are normal.

Solution 9.14

Set up the Hypothesis Test:

Since the problem is about a mean, this is a **test of a single population mean**.

$$H_0: \mu = 16.43 \quad H_a: \mu < 16.43$$

For Jeffrey to swim faster, his time will be less than 16.43 seconds. The "<" tells you this is left-tailed.

Determine the distribution needed:

Random variable: \bar{X} = the mean time to swim the 25-yard freestyle.

Distribution for the test: \bar{X} is normal (population **standard deviation** is known: $\sigma = 0.8$)

$$\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right) \text{ Therefore, } \bar{X} \sim N\left(16.43, \frac{0.8}{\sqrt{15}}\right)$$

$\mu = 16.43$ comes from H_0 and not the data. $\sigma = 0.8$, and $n = 15$.

Calculate the p -value using the normal distribution for a mean:

$$p\text{-value} = P(\bar{x} < 16) = 0.0187 \text{ where the sample mean in the problem is given as 16.}$$

$p\text{-value} = 0.0187$ (This is called the **actual level of significance**.) The p -value is the area to the left of the sample mean is given as 16.

Graph:

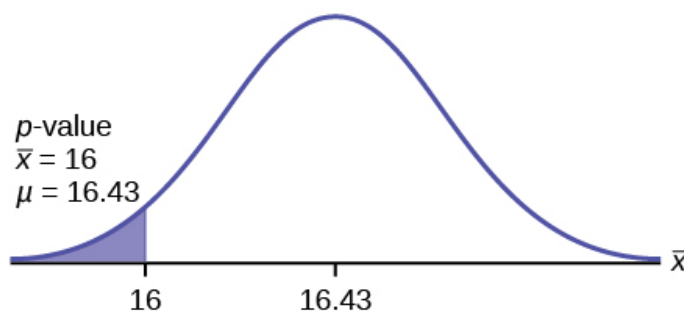


Figure 9.6

$\mu = 16.43$ comes from H_0 . Our assumption is $\mu = 16.43$.

Interpretation of the p -value: If H_0 is true, there is a 0.0187 probability (1.87%) that Jeffrey's mean time to swim the 25-yard freestyle is 16 seconds or less. Because a 1.87% chance is small, the mean time of 16 seconds or less is unlikely to have happened randomly. It is a rare event.

Compare α and the p -value:

$$\alpha = 0.05 \quad p\text{-value} = 0.0187 \quad \alpha > p\text{-value}$$

Make a decision: Since $\alpha > p\text{-value}$, reject H_0 .

This means that you reject $\mu = 16.43$. In other words, you do not think Jeffrey swims the 25-yard freestyle in 16.43 seconds but faster with the new goggles.

Conclusion: At the 5% significance level, we conclude that Jeffrey swims faster using the new goggles. The sample data show there is sufficient evidence that Jeffrey's mean time to swim the 25-yard freestyle is less than 16.43 seconds.

The p -value can easily be calculated.



Using the TI-83, 83+, 84, 84+ Calculator

Press **STAT** and arrow over to **TESTS**. Press **1:Z-Test**. Arrow over to **Stats** and press **ENTER**. Arrow down and enter 16.43 for μ_0 (null hypothesis), .8 for σ , 16 for the sample mean, and 15 for n . Arrow down to μ : (alternate hypothesis) and arrow over to $< \mu_0$. Press **ENTER**. Arrow down to **Calculate** and press **ENTER**. The calculator not only calculates the p -value ($p = 0.0187$) but it also calculates the test statistic (z -score) for the sample mean. $\mu < 16.43$ is the alternative hypothesis. Do this set of instructions again except arrow to **Draw** (instead of **Calculate**). Press **ENTER**. A shaded graph appears with $z = -2.08$ (test statistic) and $p = 0.0187$ (p -value). Make sure when you use **Draw** that no other equations are highlighted in $Y =$ and the plots are turned off.

When the calculator does a **Z-Test**, the **Z-Test** function finds the p -value by doing a normal probability calculation using the **central limit theorem**:

$$P(\bar{x} < 16) = 2\text{nd DISTR normcdf}(-10^{99}, 16, 16.43, 0.8 / \sqrt{15}).$$

The Type I and Type II errors for this problem are as follows:

The Type I error is to conclude that Jeffrey swims the 25-yard freestyle, on average, in less than 16.43 seconds when, in fact, he actually swims the 25-yard freestyle, on average, in 16.43 seconds. (Reject the null hypothesis when the null hypothesis is true.)

The Type II error is that there is not evidence to conclude that Jeffrey swims the 25-yard free-style, on average, in less than 16.43 seconds when, in fact, he actually does swim the 25-yard free-style, on average, in less than 16.43 seconds. (Do not reject the null hypothesis when the null hypothesis is false.)

Try It

9.14 The mean throwing distance of a football for a Marco, a high school freshman quarterback, is 40 yards, with a standard deviation of two yards. The team coach tells Marco to adjust his grip to get more distance. The coach records the distances for 20 throws. For the 20 throws, Marco's mean distance was 45 yards. The coach thought the different grip helped Marco throw farther than 40 yards. Conduct a hypothesis test using a preset $\alpha = 0.05$. Assume the throw distances for footballs are normal.

First, determine what type of test this is, set up the hypothesis test, find the p -value, sketch the graph, and state your conclusion.



Using the TI-83, 83+, 84, 84+ Calculator

Press **STAT** and arrow over to **TESTS**. Press **1:Z-Test**. Arrow over to **Stats** and press **ENTER**. Arrow down and enter 40 for μ_0 (null hypothesis), 2 for σ , 45 for the sample mean, and 20 for n . Arrow down to μ : (alternative hypothesis) and set it either as $<$, \neq , or $>$. Press **ENTER**. Arrow down to **Calculate** and press **ENTER**. The calculator not only calculates the p -value but it also calculates the test statistic (z -score) for the sample mean. Select $<$, \neq , or $>$ for the alternative hypothesis. Do this set of instructions again except arrow to **Draw** (instead of

Calculate). Press ENTER. A shaded graph appears with test statistic and p -value. Make sure when you use Draw that no other equations are highlighted in $Y =$ and the plots are turned off.

HISTORICAL NOTE (EXAMPLE 9.11)

The traditional way to compare the two probabilities, α and the p -value, is to compare the critical value (z-score from α) to the test statistic (z-score from data). The calculated test statistic for the p -value is -2.08 . (From the Central Limit Theorem, the test statistic formula is $z = \frac{\bar{x} - \mu_X}{(\frac{\sigma_X}{\sqrt{n}})}$. For this problem, $\bar{x} = 16$, $\mu_X = 16.43$ from the null hypotheses

is, $\sigma_X = 0.8$, and $n = 15$.) You can find the critical value for $\alpha = 0.05$ in the normal table (see **15.Tables** in the Table of Contents). The z-score for an area to the left equal to 0.05 is midway between -1.65 and -1.64 (0.05 is midway between 0.0505 and 0.0495). The z-score is -1.645 . Since $-1.645 > -2.08$ (which demonstrates that $\alpha > p$ -value), reject H_0 . Traditionally, the decision to reject or not reject was done in this way. Today, comparing the two probabilities α and the p -value is very common. For this problem, the p -value, 0.0187 is considerably smaller than α , 0.05. You can be confident about your decision to reject. The graph shows α , the p -value, and the test statistics and the critical value.

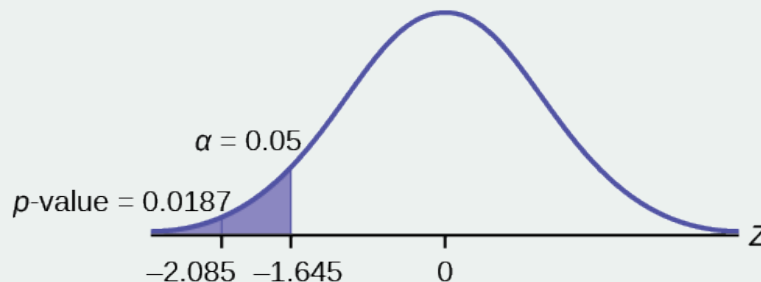


Figure 9.7

Example 9.15

A college football coach thought that his players could bench press a **mean weight of 275 pounds**. It is known that the **standard deviation is 55 pounds**. Three of his players thought that the mean weight was **more than** that amount. They asked **30** of their teammates for their estimated maximum lift on the bench press exercise. The data ranged from 205 pounds to 385 pounds. The actual different weights were (frequencies are in parentheses) 205(3); 215(3); 225(1); 241(2); 252(2); 265(2); 275(2); 313(2); 316(5); 338(2); 341(1); 345(2); 368(2); 385(1).

Conduct a hypothesis test using a 2.5% level of significance to determine if the bench press mean is **more than 275 pounds**.

Solution 9.15

Set up the Hypothesis Test:

Since the problem is about a mean weight, this is a **test of a single population mean**.

$$H_0: \mu = 275$$

$$H_a: \mu > 275$$

This is a right-tailed test.

Calculating the distribution needed:

Random variable: \bar{X} = the mean weight, in pounds, lifted by the football players.

Distribution for the test: It is normal because σ is known.

$$\bar{X} \sim N\left(275, \frac{55}{\sqrt{30}}\right)$$

$\bar{x} = 286.2$ pounds (from the data).

$\sigma = 55$ pounds (**Always use σ if you know it.**) We assume $\mu = 275$ pounds unless our data shows us otherwise.

Calculate the p -value using the normal distribution for a mean and using the sample mean as input (see **Appendix G** for using the data as input):

$$p\text{-value} = P(\bar{x} > 286.2) = 0.1323.$$

Interpretation of the p -value: If H_0 is true, then there is a 0.1331 probability (13.23%) that the football players can lift a mean weight of 286.2 pounds or more. Because a 13.23% chance is large enough, a mean weight lift of 286.2 pounds or more is not a rare event.

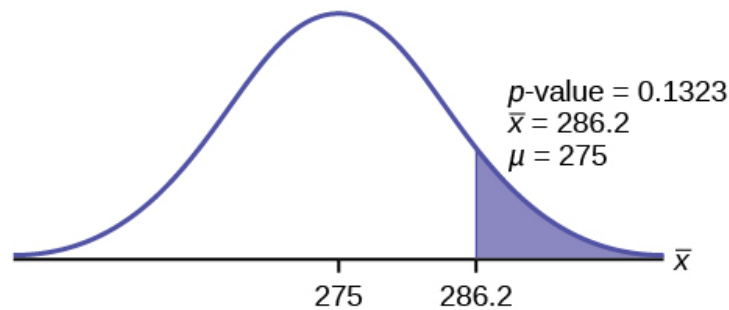


Figure 9.8

Compare α and the p -value:

$$\alpha = 0.025 \quad p\text{-value} = 0.1323$$

Make a decision: Since $\alpha < p\text{-value}$, do not reject H_0 .

Conclusion: At the 2.5% level of significance, from the sample data, there is not sufficient evidence to conclude that the true mean weight lifted is more than 275 pounds.

The p -value can easily be calculated.



Using the TI-83, 83+, 84, 84+ Calculator

Put the data and frequencies into lists. Press **STAT** and arrow over to **TESTS**. Press **1:Z-Test**. Arrow over to **Data** and press **ENTER**. Arrow down and enter 275 for μ_0 , 55 for σ , the name of the list where you put the data, and the name of the list where you put the frequencies. Arrow down to μ : and arrow over to $> \mu_0$. Press **ENTER**. Arrow down to **Calculate** and press **ENTER**. The calculator not only calculates the p -value ($p = 0.1331$, a little different from the previous calculation - in it we used the sample mean rounded to one decimal place instead of the data) but it also calculates the test statistic (z -score) for the sample mean, the sample mean, and the sample standard deviation. $\mu > 275$ is the alternative hypothesis. Do this set of instructions again except arrow to **Draw** (instead of **Calculate**). Press **ENTER**. A shaded graph appears with $z = 1.112$ (test statistic) and $p = 0.1331$ (p -value). Make sure when you use **Draw** that no other equations are highlighted in $Y =$ and the plots are turned off.

Example 9.16

Statistics students believe that the mean score on the first statistics test is 65. A statistics instructor thinks the mean score is higher than 65. He samples ten statistics students and obtains the scores 65; 65; 70; 67; 66; 63; 63; 68; 72; 71. He performs a hypothesis test using a 5% level of significance. The data are assumed to be from a normal distribution.

Solution 9.16

Set up the hypothesis test:

A 5% level of significance means that $\alpha = 0.05$. This is a test of a **single population mean**.

$$H_0: \mu = 65 \quad H_a: \mu > 65$$

Since the instructor thinks the average score is higher, use a ">". The ">" means the test is right-tailed.

Determine the distribution needed:

Random variable: \bar{X} = average score on the first statistics test.

Distribution for the test: If you read the problem carefully, you will notice that there is **no population standard deviation given**. You are only given $n = 10$ sample data values. Notice also that the data come from a normal distribution. This means that the distribution for the test is a student's t .

Use t_{df} . Therefore, the distribution for the test is t_9 where $n = 10$ and $df = 10 - 1 = 9$.

Calculate the p -value using the Student's t -distribution:

$p\text{-value} = P(\bar{x} > 67) = 0.0396$ where the sample mean and sample standard deviation are calculated as 67 and 3.1972 from the data.

Interpretation of the p -value: If the null hypothesis is true, then there is a 0.0396 probability (3.96%) that the sample mean is 65 or more.

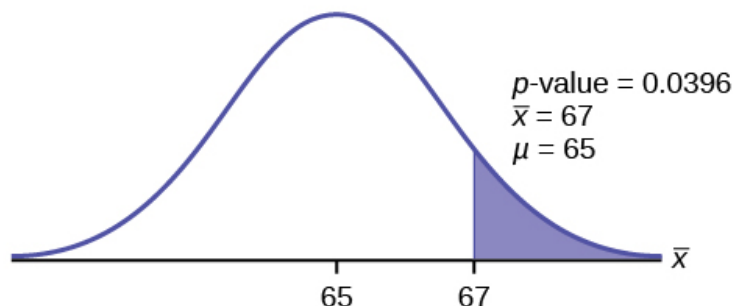


Figure 9.9

Compare α and the p -value:

Since $\alpha = 0.05$ and $p\text{-value} = 0.0396$, $\alpha > p\text{-value}$.

Make a decision: Since $\alpha > p\text{-value}$, reject H_0 .

This means you reject $\mu = 65$. In other words, you believe the average test score is more than 65.

Conclusion: At a 5% level of significance, the sample data show sufficient evidence that the mean (average) test score is more than 65, just as the math instructor thinks.

The p -value can easily be calculated.



Using the TI-83, 83+, 84, 84+ Calculator

Put the data into a list. Press **STAT** and arrow over to **TESTS**. Press **2:T-Test**. Arrow over to **Data** and press **ENTER**. Arrow down and enter 65 for μ_0 , the name of the list where you put the data, and 1 for **Freq**:. Arrow down to μ : and arrow over to $> \mu_0$. Press **ENTER**. Arrow down to **Calculate** and press **ENTER**. The calculator not only calculates the p -value ($p = 0.0396$) but it also calculates the test statistic (t -score) for the sample mean, the sample mean, and the sample standard deviation. $\mu > 65$ is the alternative hypothesis. Do this set of instructions again except arrow to **Draw** (instead of **Calculate**). Press **ENTER**. A shaded graph appears with $t = 1.9781$ (test statistic) and $p = 0.0396$ (p -value). Make sure when you use **Draw** that no other equations are highlighted in $Y =$ and the plots are turned off.

Try It

9.16 It is believed that a stock price for a particular company will grow at a rate of \$5 per week with a standard deviation of \$1. An investor believes the stock won't grow as quickly. The changes in stock price is recorded for ten weeks and are as follows: \$4, \$3, \$2, \$3, \$1, \$7, \$2, \$1, \$1, \$2. Perform a hypothesis test using a 5% level of significance. State the null and alternative hypotheses, find the p -value, state your conclusion, and identify the Type I and Type II errors.

Example 9.17

Joon believes that 50% of first-time brides in the United States are younger than their grooms. She performs a hypothesis test to determine if the percentage is **the same or different from 50%**. Joon samples **100 first-time brides** and 53 reply that they are younger than their grooms. For the hypothesis test, she uses a 1% level of significance.

Solution 9.17

Set up the hypothesis test:

The 1% level of significance means that $\alpha = 0.01$. This is a **test of a single population proportion**.

$$H_0: p = 0.50 \quad H_a: p \neq 0.50$$

The words "**is the same or different from**" tell you this is a two-tailed test.

Calculate the distribution needed:

Random variable: P' = the percent of first-time brides who are younger than their grooms.

Distribution for the test: The problem contains no mention of a mean. The information is given in terms of percentages. Use the distribution for P' , the estimated proportion.

$$P' \sim N\left(p, \sqrt{\frac{p \cdot q}{n}}\right) \text{ Therefore, } P' \sim N\left(0.5, \sqrt{\frac{0.5 \cdot 0.5}{100}}\right)$$

where $p = 0.50$, $q = 1 - p = 0.50$, and $n = 100$

Calculate the p -value using the normal distribution for proportions:

$$p\text{-value} = P(p' < 0.47 \text{ or } p' > 0.53) = 0.5485$$

$$\text{where } x = 53, p' = \frac{x}{n} = \frac{53}{100} = 0.53.$$

Interpretation of the p -value: If the null hypothesis is true, there is 0.5485 probability (54.85%) that the sample (estimated) proportion p' is 0.53 or more OR 0.47 or less (see the graph in **Figure 9.9**).

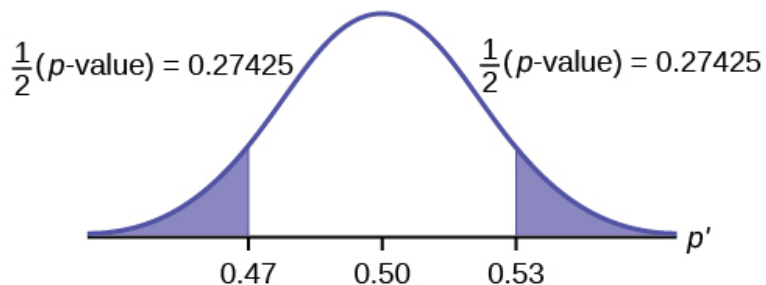


Figure 9.10

$\mu = p = 0.50$ comes from H_0 , the null hypothesis.

$p' = 0.53$. Since the curve is symmetrical and the test is two-tailed, the p' for the left tail is equal to $0.50 - 0.03 = 0.47$ where $\mu = p = 0.50$. (0.03 is the difference between 0.53 and 0.50.)

Compare α and the p -value:

Since $\alpha = 0.01$ and $p\text{-value} = 0.5485$, $\alpha < p\text{-value}$.

Make a decision: Since $\alpha < p\text{-value}$, you cannot reject H_0 .

Conclusion: At the 1% level of significance, the sample data do not show sufficient evidence that the percentage of first-time brides who are younger than their grooms is different from 50%.

The p -value can easily be calculated.



Using the TI-83, 83+, 84, 84+ Calculator

Press STAT and arrow over to TESTS. Press 5:1-PropZTest. Enter .5 for p_0 , 53 for x and 100 for n . Arrow down to Prop and arrow to not equals p_0 . Press ENTER. Arrow down to Calculate and press ENTER. The calculator calculates the p -value ($p = 0.5485$) and the test statistic (z -score). Prop not equals .5 is the alternate hypothesis. Do this set of instructions again except arrow to Draw (instead of Calculate). Press ENTER. A shaded graph appears with $z = 0.6$ (test statistic) and $p = 0.5485$ (p -value). Make sure when you use Draw that no other equations are highlighted in $Y =$ and the plots are turned off.

The Type I and Type II errors are as follows:

The Type I error is to conclude that the proportion of first-time brides who are younger than their grooms is different from 50% when, in fact, the proportion is actually 50%. (Reject the null hypothesis when the null hypothesis is true).

The Type II error is there is not enough evidence to conclude that the proportion of first time brides who are younger than their grooms differs from 50% when, in fact, the proportion does differ from 50%. (Do not reject the null hypothesis when the null hypothesis is false.)

Try It Σ

9.17 A teacher believes that 85% of students in the class will want to go on a field trip to the local zoo. She performs a hypothesis test to determine if the percentage is the same or different from 85%. The teacher samples 50 students and 39 reply that they would want to go to the zoo. For the hypothesis test, use a 1% level of significance.

First, determine what type of test this is, set up the hypothesis test, find the p -value, sketch the graph, and state your conclusion.

Example 9.18

Suppose a consumer group suspects that the proportion of households that have three cell phones is 30%. A cell phone company has reason to believe that the proportion is not 30%. Before they start a big advertising campaign, they conduct a hypothesis test. Their marketing people survey 150 households with the result that 43 of the households have three cell phones.

Solution 9.18

Set up the Hypothesis Test:

$$H_0: p = 0.30 \quad H_a: p \neq 0.30$$

Determine the distribution needed:

The **random variable** is $P' =$ proportion of households that have three cell phones.

The **distribution** for the hypothesis test is $P' \sim N\left(0.30, \sqrt{\frac{(0.30) \cdot (0.70)}{150}}\right)$

a. The value that helps determine the p -value is p' . Calculate p' .

Solution 9.18

a. $p' = \frac{x}{n}$ where x is the number of successes and n is the total number in the sample.

$$x = 43, n = 150$$

$$p' = \frac{43}{150}$$

b. What is a **success** for this problem?

Solution 9.18

b. A success is having three cell phones in a household.

c. What is the level of significance?

Solution 9.18

c. The level of significance is the preset α . Since α is not given, assume that $\alpha = 0.05$.

d. Draw the graph for this problem. Draw the horizontal axis. Label and shade appropriately. Calculate the p -value.

Solution 9.18

d. p -value = 0.7216

e. Make a decision. _____ (Reject/Do not reject) H_0 because _____.

Solution 9.18

e. Assuming that $\alpha = 0.05$, $\alpha < p$ -value. The decision is do not reject H_0 because there is not sufficient evidence to conclude that the proportion of households that have three cell phones is not 30%.

Try It

9.18 Marketers believe that 92% of adults in the United States own a cell phone. A cell phone manufacturer believes that number is actually lower. 200 American adults are surveyed, of which, 174 report having cell phones. Use a 5%

level of significance. State the null and alternative hypothesis, find the p -value, state your conclusion, and identify the Type I and Type II errors.

The next example is a poem written by a statistics student named Nicole Hart. The solution to the problem follows the poem. Notice that the hypothesis test is for a single population proportion. This means that the null and alternate hypotheses use the parameter p . The distribution for the test is normal. The estimated proportion p' is the proportion of fleas killed to the total fleas found on Fido. This is sample information. The problem gives a preconceived $\alpha = 0.01$, for comparison, and a 95% confidence interval computation. The poem is clever and humorous, so please enjoy it!

Example 9.19

My dog has so many fleas,
They do not come off with ease.
As for shampoo, I have tried many types
Even one called Bubble Hype,
Which only killed 25% of the fleas,
Unfortunately I was not pleased.

I've used all kinds of soap,
Until I had given up hope
Until one day I saw
An ad that put me in awe.

A shampoo used for dogs
Called GOOD ENOUGH to Clean a Hog
Guaranteed to kill more fleas.

I gave Fido a bath
And after doing the math
His number of fleas
Started dropping by 3's!

Before his shampoo
I counted 42.
At the end of his bath,
I redid the math
And the new shampoo had killed 17 fleas.
So now I was pleased.

Now it is time for you to have some fun
With the level of significance being .01,
You must help me figure out
Use the new shampoo or go without?

Solution 9.19

Set up the hypothesis test:

$$H_0: p \leq 0.25 \quad H_a: p > 0.25$$

Determine the distribution needed:

In words, CLEARLY state what your random variable \bar{X} or P' represents.

P' = The proportion of fleas that are killed by the new shampoo

State the distribution to use for the test.

Normal: $N\left(0.25, \sqrt{\frac{(0.25)(1 - 0.25)}{42}}\right)$

Test Statistic: $z = 2.3163$

Calculate the p -value using the normal distribution for proportions:

$p\text{-value} = 0.0103$

In one to two complete sentences, explain what the p -value means for this problem.

If the null hypothesis is true (the proportion is 0.25), then there is a 0.0103 probability that the sample (estimated) proportion is 0.4048 ($\frac{17}{42}$) or more.

Use the previous information to sketch a picture of this situation. CLEARLY, label and scale the horizontal axis and shade the region(s) corresponding to the p -value.

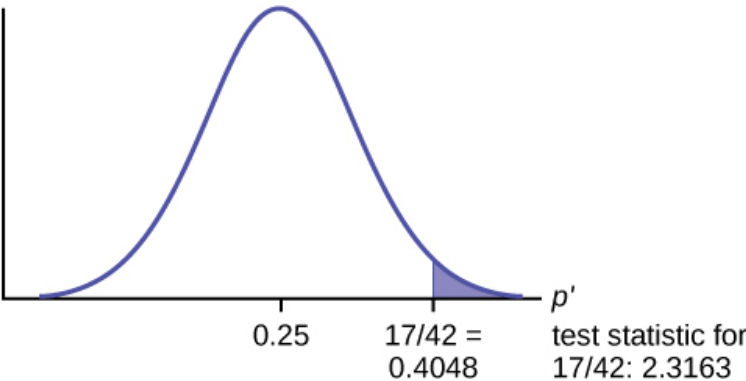


Figure 9.11

Compare α and the p -value:

Indicate the correct decision (“reject” or “do not reject” the null hypothesis), the reason for it, and write an appropriate conclusion, using complete sentences.

alpha	decision	reason for decision
0.01	Do not reject H_0	$\alpha < p\text{-value}$

Table 9.3

Conclusion: At the 1% level of significance, the sample data do not show sufficient evidence that the percentage of fleas that are killed by the new shampoo is more than 25%.

Construct a 95% confidence interval for the true mean or proportion. Include a sketch of the graph of the situation. Label the point estimate and the lower and upper bounds of the confidence interval.

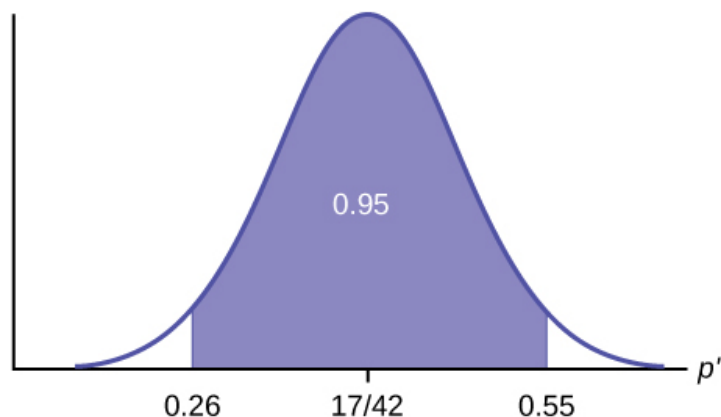


Figure 9.12

Confidence Interval: (0.26,0.55) We are 95% confident that the true population proportion p of fleas that are killed by the new shampoo is between 26% and 55%.

NOTE

This test result is not very definitive since the p -value is very close to alpha. In reality, one would probably do more tests by giving the dog another bath after the fleas have had a chance to return.

Example 9.20

The National Institute of Standards and Technology provides exact data on conductivity properties of materials. Following are conductivity measurements for 11 randomly selected pieces of a particular type of glass.

1.11; 1.07; 1.11; 1.07; 1.12; 1.08; .98; .98 1.02; .95; .95

Is there convincing evidence that the average conductivity of this type of glass is greater than one? Use a significance level of 0.05. Assume the population is normal.

Solution 9.20

Let's follow a four-step process to answer this statistical question.

1. **State the Question:** We need to determine if, at a 0.05 significance level, the average conductivity of the selected glass is greater than one. Our hypotheses will be
 - a. $H_0: \mu \leq 1$
 - b. $H_a: \mu > 1$
2. **Plan:** We are testing a sample mean without a known population standard deviation. Therefore, we need to use a Student's-t distribution. Assume the underlying population is normal.
3. **Do the calculations:** We will input the sample data into the TI-83 as follows.

L1	L2	L3	1
1.08			
.98			
.98			
1.02			
.95			
.95			
77777			
L1(12) =			

Figure 9.13

T-Test	
Inpt:	Stats
μ_0 :	1
List:	L1
Freq:	1
μ :	$\neq \mu_0$ $< \mu_0$ $> \mu_0$
Calculate	Draw

Figure 9.14

T-Test	
μ :	> 1
t:	2.01377743
P:	.035860646
\bar{x} :	1.04
Sx:	.0658786764
n:	11

Figure 9.15

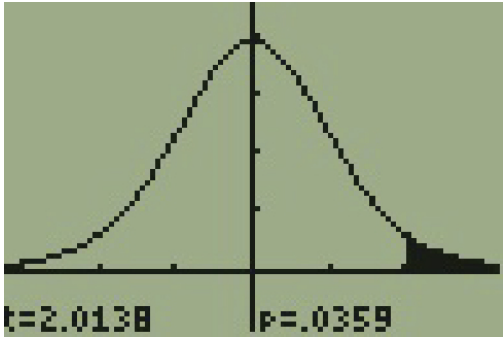


Figure 9.16

4. **State the Conclusions:** Since the p -value* ($p = 0.036$) is less than our alpha value, we will reject the null hypothesis. It is reasonable to state that the data supports the claim that the average conductivity level is greater than one.

Example 9.21

In a study of 420,019 cell phone users, 172 of the subjects developed brain cancer. Test the claim that cell phone users developed brain cancer at a greater rate than that for non-cell phone users (the rate of brain cancer for non-cell phone users is 0.0340%). Since this is a critical issue, use a 0.005 significance level. Explain why the significance level should be so low in terms of a Type I error.

Solution 9.21

We will follow the four-step process.

1. We need to conduct a hypothesis test on the claimed cancer rate. Our hypotheses will be
 - a. $H_0: p \leq 0.00034$
 - b. $H_a: p > 0.00034$

If we commit a Type I error, we are essentially accepting a false claim. Since the claim describes cancer-causing environments, we want to minimize the chances of incorrectly identifying causes of cancer.

2. We will be testing a sample proportion with $x = 172$ and $n = 420,019$. The sample is sufficiently large because we have $np = 420,019(0.00034) = 142.8$, $nq = 420,019(0.99966) = 419,876.2$, two independent outcomes, and a fixed probability of success $p = 0.00034$. Thus we will be able to generalize our results to the population.
3. The associated TI results are

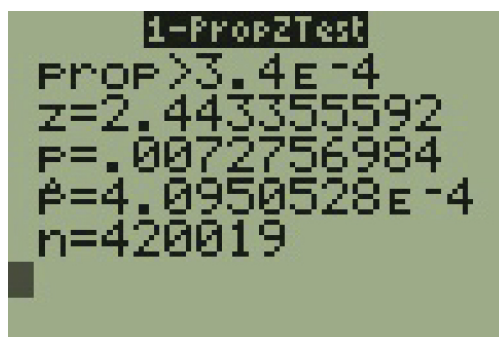


Figure 9.17

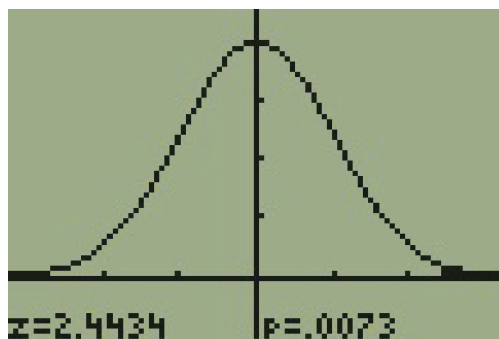


Figure 9.18

4. Since the p -value = 0.0073 is greater than our alpha value = 0.005, we cannot reject the null. Therefore, we conclude that there is not enough evidence to support the claim of higher brain cancer rates for the cell phone users.

Example 9.22

According to the US Census there are approximately 268,608,618 residents aged 12 and older. Statistics from the Rape, Abuse, and Incest National Network indicate that, on average, 207,754 rapes occur each year (male and female) for persons aged 12 and older. This translates into a percentage of sexual assaults of 0.078%. In Daviess County, KY, there were reported 11 rapes for a population of 37,937. Conduct an appropriate hypothesis test to determine if there is a statistically significant difference between the local sexual assault percentage and the national sexual assault percentage. Use a significance level of 0.01.

Solution 9.22

We will follow the four-step plan.

1. We need to test whether the proportion of sexual assaults in Daviess County, KY is significantly different from the national average.
2. Since we are presented with proportions, we will use a one-proportion z -test. The hypotheses for the test will be
 - a. $H_0: p = 0.00078$
 - b. $H_a: p \neq 0.00078$
3. The following screen shots display the summary statistics from the hypothesis test.



Figure 9.19

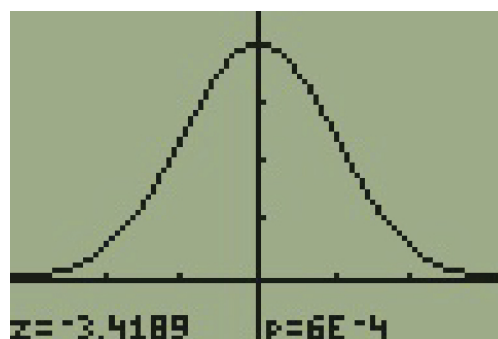


Figure 9.20

4. Since the p -value, $p = 0.00063$, is less than the alpha level of 0.01, the sample data indicates that we should reject the null hypothesis. In conclusion, the sample data support the claim that the proportion of sexual assaults in Daviess County, Kentucky is different from the national average proportion.

9.6 | Hypothesis Testing of a Single Mean and Single Proportion

9.1 Hypothesis Testing of a Single Mean and Single Proportion

Class Time:

Names:

Student Learning Outcomes

- The student will select the appropriate distributions to use in each case.
- The student will conduct hypothesis tests and interpret the results.

Television Survey

In a recent survey, it was stated that Americans watch television on average four hours per day. Assume that $\sigma = 2$. Using your class as the sample, conduct a hypothesis test to determine if the average for students at your school is lower.

1. H_0 : _____
2. H_a : _____
3. In words, define the random variable. _____ = _____
4. The distribution to use for the test is _____.
5. Determine the test statistic using your data.
6. Draw a graph and label it appropriately. Shade the actual level of significance.
 - a. Graph:

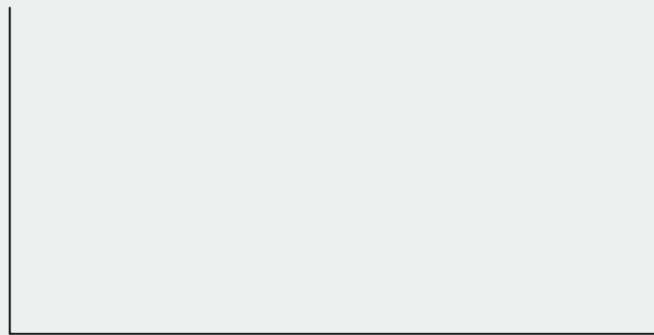


Figure 9.21

- b. Determine the p -value.
7. Do you or do you not reject the null hypothesis? Why?
 8. Write a clear conclusion using a complete sentence.

Language Survey

About 42.3% of Californians and 19.6% of all Americans over age five speak a language other than English at home. Using your class as the sample, conduct a hypothesis test to determine if the percent of the students at your school who speak a language other than English at home is different from 42.3%.

1. H_0 : _____
2. H_a : _____

3. In words, define the random variable. _____ = _____
4. The distribution to use for the test is _____
5. Determine the test statistic using your data.
6. Draw a graph and label it appropriately. Shade the actual level of significance.
 - a. Graph:

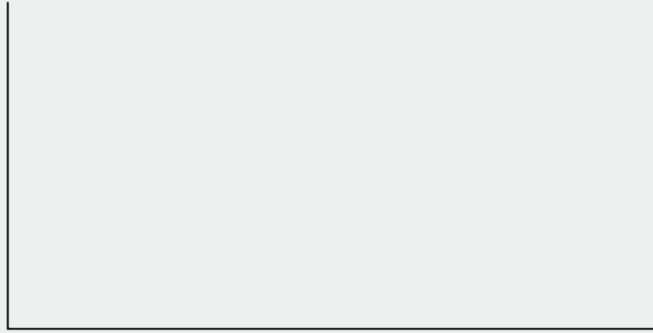


Figure 9.22

- b. Determine the p -value.
7. Do you or do you not reject the null hypothesis? Why?
8. Write a clear conclusion using a complete sentence.

Jeans Survey

Suppose that young adults own an average of three pairs of jeans. Survey eight people from your class to determine if the average is higher than three. Assume the population is normal.

1. H_0 : _____
2. H_a : _____
3. In words, define the random variable. _____ = _____
4. The distribution to use for the test is _____.
5. Determine the test statistic using your data.
6. Draw a graph and label it appropriately. Shade the actual level of significance.
 - a. Graph:

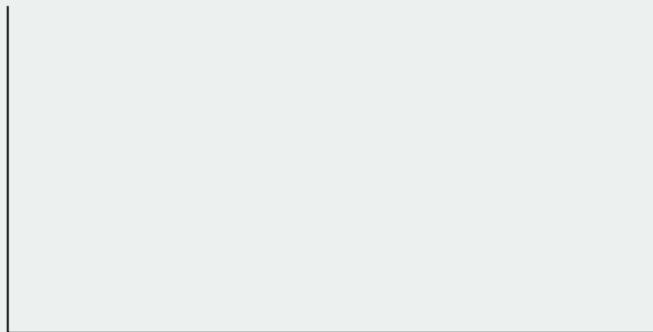


Figure 9.23

- b. Determine the p -value.
7. Do you or do you not reject the null hypothesis? Why?

8. Write a clear conclusion using a complete sentence.

KEY TERMS

Binomial Distribution a discrete random variable (RV) that arises from Bernoulli trials. There are a fixed number, n , of independent trials. “Independent” means that the result of any trial (for example, trial 1) does not affect the results of the following trials, and all trials are conducted under the same conditions. Under these circumstances the binomial RV X is defined as the number of successes in n trials. The notation is: $X \sim B(n, p)$ $\mu = np$ and the standard deviation is $\sigma = \sqrt{npq}$. The probability of exactly x successes in n trials is $P(X = x) = \binom{n}{x} p^x q^{n-x}$.

Central Limit Theorem Given a random variable (RV) with known mean μ and known standard deviation σ . We are sampling with size n and we are interested in two new RVs - the sample mean, \bar{X} , and the sample sum, ΣX . If the size n of the sample is sufficiently large, then $\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$ and $\Sigma X \sim N(n\mu, \sqrt{n}\sigma)$. If the size n of the sample is sufficiently large, then the distribution of the sample means and the distribution of the sample sums will approximate a normal distribution regardless of the shape of the population. The mean of the sample means will equal the population mean and the mean of the sample sums will equal n times the population mean. The standard deviation of the distribution of the sample means, $\frac{\sigma}{\sqrt{n}}$, is called the standard error of the mean.

Confidence Interval (CI) an interval estimate for an unknown population parameter. This depends on:

- The desired confidence level.
- Information that is known about the distribution (for example, known standard deviation).
- The sample and its size.

Hypothesis a statement about the value of a population parameter, in case of two hypotheses, the statement assumed to be true is called the null hypothesis (notation H_0) and the contradictory statement is called the alternative hypothesis (notation H_a).

Hypothesis Testing Based on sample evidence, a procedure for determining whether the hypothesis stated is a reasonable statement and should not be rejected, or is unreasonable and should be rejected.

Level of Significance of the Test probability of a Type I error (reject the null hypothesis when it is true). Notation: α . In hypothesis testing, the Level of Significance is called the preconceived α or the preset α .

Normal Distribution

a continuous random variable (RV) with pdf $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$, where μ is the mean of

the distribution, and σ is the standard deviation, notation: $X \sim N(\mu, \sigma)$. If $\mu = 0$ and $\sigma = 1$, the RV is called **the standard normal distribution**.

p-value the probability that an event will happen purely by chance assuming the null hypothesis is true. The smaller the p-value, the stronger the evidence is against the null hypothesis.

Standard Deviation a number that is equal to the square root of the variance and measures how far data values are from their mean; notation: s for sample standard deviation and σ for population standard deviation.

Student's t-Distribution investigated and reported by William S. Gossett in 1908 and published under the pseudonym Student. The major characteristics of the random variable (RV) are:

- It is continuous and assumes any real values.
- The pdf is symmetrical about its mean of zero. However, it is more spread out and flatter at the apex than the normal distribution.
- It approaches the standard normal distribution as n gets larger.
- There is a "family" of t distributions: every representative of the family is completely defined by the number of degrees of freedom which is one less than the number of data items.

Type 1 Error The decision is to reject the null hypothesis when, in fact, the null hypothesis is true.

Type 2 Error The decision is not to reject the null hypothesis when, in fact, the null hypothesis is false.

CHAPTER REVIEW

9.1 Null and Alternative Hypotheses

In a **hypothesis test**, sample data is evaluated in order to arrive at a decision about some type of claim. If certain conditions about the sample are satisfied, then the claim can be evaluated for a population. In a hypothesis test, we:

1. Evaluate the **null hypothesis**, typically denoted with H_0 . The null is not rejected unless the hypothesis test shows otherwise. The null statement must always contain some form of equality ($=$, \leq or \geq)
2. Always write the **alternative hypothesis**, typically denoted with H_a or H_1 , using less than, greater than, or not equals symbols, i.e., (\neq , $>$, or $<$).
3. If we reject the null hypothesis, then we can assume there is enough evidence to support the alternative hypothesis.
4. Never state that a claim is proven true or false. Keep in mind the underlying fact that hypothesis testing is based on probability laws; therefore, we can talk only in terms of non-absolute certainties.

9.2 Outcomes and the Type I and Type II Errors

In every hypothesis test, the outcomes are dependent on a correct interpretation of the data. Incorrect calculations or misunderstood summary statistics can yield errors that affect the results. A **Type I** error occurs when a true null hypothesis is rejected. A **Type II error** occurs when a false null hypothesis is not rejected.

The probabilities of these errors are denoted by the Greek letters α and β , for a Type I and a Type II error respectively. The power of the test, $1 - \beta$, quantifies the likelihood that a test will yield the correct result of a true alternative hypothesis being accepted. A high power is desirable.

9.3 Distribution Needed for Hypothesis Testing

In order for a hypothesis test's results to be generalized to a population, certain requirements must be satisfied.

When testing for a single population mean:

1. A Student's t -test should be used if the data come from a simple, random sample and the population is approximately normally distributed, or the sample size is large, with an unknown standard deviation.
2. The normal test will work if the data come from a simple, random sample and the population is approximately normally distributed, or the sample size is large, with a known standard deviation.

When testing a single population proportion use a normal test for a single population proportion if the data comes from a simple, random sample, fill the requirements for a binomial distribution, and the mean number of success and the mean number of failures satisfy the conditions: $np > 5$ and $nq > 5$ where n is the sample size, p is the probability of a success, and q is the probability of a failure.

9.4 Rare Events, the Sample, Decision and Conclusion

When the probability of an event occurring is low, and it happens, it is called a rare event. Rare events are important to consider in hypothesis testing because they can inform your willingness not to reject or to reject a null hypothesis. To test a null hypothesis, find the p -value for the sample data and graph the results. When deciding whether or not to reject the null hypothesis, keep these two parameters in mind:

1. $\alpha > p$ -value, reject the null hypothesis
2. $\alpha \leq p$ -value, do not reject the null hypothesis

9.5 Additional Information and Full Hypothesis Test Examples

The **hypothesis test** itself has an established process. This can be summarized as follows:

1. Determine H_0 and H_a . Remember, they are contradictory.
2. Determine the random variable.
3. Determine the distribution for the test.

4. Draw a graph, calculate the test statistic, and use the test statistic to calculate the p -value. (A z -score and a t -score are examples of test statistics.)
5. Compare the preconceived α with the p -value, make a decision (reject or do not reject H_0), and write a clear conclusion using English sentences.

Notice that in performing the hypothesis test, you use α and not β . β is needed to help determine the sample size of the data that is used in calculating the p -value. Remember that the quantity $1 - \beta$ is called the **Power of the Test**. A high power is desirable. If the power is too low, statisticians typically increase the sample size while keeping α the same. If the power is low, the null hypothesis might not be rejected when it should be.

FORMULA REVIEW

9.1 Null and Alternative Hypotheses

H_0 and H_a are contradictory.

If H_0 has:	equal (=)	greater than or equal to (\geq)	less than or equal to (\leq)
then H_a has:	not equal (\neq) or greater than ($>$) or less than ($<$)	less than ($<$)	greater than ($>$)

Table 9.4

If $\alpha \leq p$ -value, then do not reject H_0 .

If $\alpha > p$ -value, then reject H_0 .

α is preconceived. Its value is set before the hypothesis test starts. The p -value is calculated from the data.

9.2 Outcomes and the Type I and Type II Errors

α = probability of a Type I error = $P(\text{Type I error})$ = probability of rejecting the null hypothesis when the null hypothesis is true.

β = probability of a Type II error = $P(\text{Type II error})$ = probability of not rejecting the null hypothesis when the null hypothesis is false.

9.3 Distribution Needed for Hypothesis Testing

If there is no given preconceived α , then use $\alpha = 0.05$.

Types of Hypothesis Tests

- Single population mean, **known** population variance (or standard deviation): **Normal test**.
- Single population mean, **unknown** population variance (or standard deviation): **Student's t -test**.
- Single population proportion: **Normal test**.
- For a **single population mean**, we may use a normal distribution with the following mean and standard deviation. Means: $\mu = \mu_x$ and $\sigma_x = \frac{\sigma_x}{\sqrt{n}}$.
- A **single population proportion**, we may use a normal distribution with the following mean and standard deviation. Proportions: $\mu = p$ and $\sigma = \sqrt{\frac{pq}{n}}$.

PRACTICE

9.1 Null and Alternative Hypotheses

1. You are testing that the mean speed of your cable Internet connection is more than three Megabits per second. What is the random variable? Describe in words.
2. You are testing that the mean speed of your cable Internet connection is more than three Megabits per second. State the null and alternative hypotheses.
3. The American family has an average of two children. What is the random variable? Describe in words.
4. The mean entry level salary of an employee at a company is \$58,000. You believe it is higher for IT professionals in the company. State the null and alternative hypotheses.
5. A sociologist claims the probability that a person picked at random in Times Square in New York City is visiting the area is 0.83. You want to test to see if the proportion is actually less. What is the random variable? Describe in words.
6. A sociologist claims the probability that a person picked at random in Times Square in New York City is visiting the area is 0.83. You want to test to see if the claim is correct. State the null and alternative hypotheses.
7. In a population of fish, approximately 42% are female. A test is conducted to see if, in fact, the proportion is less. State the null and alternative hypotheses.
8. Suppose that a recent article stated that the mean time spent in jail by a first-time convicted burglar is 2.5 years. A study was then done to see if the mean time has increased in the new century. A random sample of 26 first-time convicted burglars

in a recent year was picked. The mean length of time in jail from the survey was 3 years with a standard deviation of 1.8 years. Suppose that it is somehow known that the population standard deviation is 1.5. If you were conducting a hypothesis test to determine if the mean length of jail time has increased, what would the null and alternative hypotheses be? The distribution of the population is normal.

- a. H_0 : _____
- b. H_a : _____

9. A random survey of 75 death row inmates revealed that the mean length of time on death row is 17.4 years with a standard deviation of 6.3 years. If you were conducting a hypothesis test to determine if the population mean time on death row could likely be 15 years, what would the null and alternative hypotheses be?

- a. H_0 : _____
- b. H_a : _____

10. The National Institute of Mental Health published an article stating that in any one-year period, approximately 9.5 percent of American adults suffer from depression or a depressive illness. Suppose that in a survey of 100 people in a certain town, seven of them suffered from depression or a depressive illness. If you were conducting a hypothesis test to determine if the true proportion of people in that town suffering from depression or a depressive illness is lower than the percent in the general adult American population, what would the null and alternative hypotheses be?

- a. H_0 : _____
- b. H_a : _____

9.2 Outcomes and the Type I and Type II Errors

11. The mean price of mid-sized cars in a region is \$32,000. A test is conducted to see if the claim is true. State the Type I and Type II errors in complete sentences.

12. A sleeping bag is tested to withstand temperatures of -15°F . You think the bag cannot stand temperatures that low. State the Type I and Type II errors in complete sentences.

13. For **Exercise 9.12**, what are α and β in words?

14. In words, describe $1 - \beta$ For **Exercise 9.12**.

15. A group of doctors is deciding whether or not to perform an operation. Suppose the null hypothesis, H_0 , is: the surgical procedure will go well. State the Type I and Type II errors in complete sentences.

16. A group of doctors is deciding whether or not to perform an operation. Suppose the null hypothesis, H_0 , is: the surgical procedure will go well. Which is the error with the greater consequence?

17. The power of a test is 0.981. What is the probability of a Type II error?

18. A group of divers is exploring an old sunken ship. Suppose the null hypothesis, H_0 , is: the sunken ship does not contain buried treasure. State the Type I and Type II errors in complete sentences.

19. A microbiologist is testing a water sample for E-coli. Suppose the null hypothesis, H_0 , is: the sample does not contain E-coli. The probability that the sample does not contain E-coli, but the microbiologist thinks it does is 0.012. The probability that the sample does contain E-coli, but the microbiologist thinks it does not is 0.002. What is the power of this test?

20. A microbiologist is testing a water sample for E-coli. Suppose the null hypothesis, H_0 , is: the sample contains E-coli. Which is the error with the greater consequence?

9.3 Distribution Needed for Hypothesis Testing

21. Which two distributions can you use for hypothesis testing for this chapter?

22. Which distribution do you use when you are testing a population mean and the standard deviation is known? Assume sample size is large.

23. Which distribution do you use when the standard deviation is not known and you are testing one population mean? Assume sample size is large.

24. A population mean is 13. The sample mean is 12.8, and the sample standard deviation is two. The sample size is 20. What distribution should you use to perform a hypothesis test? Assume the underlying population is normal.

25. A population has a mean is 25 and a standard deviation of five. The sample mean is 24, and the sample size is 108. What distribution should you use to perform a hypothesis test?

26. It is thought that 42% of respondents in a taste test would prefer Brand A. In a particular test of 100 people, 39% preferred Brand A. What distribution should you use to perform a hypothesis test?

27. You are performing a hypothesis test of a single population mean using a Student's t -distribution. What must you assume about the distribution of the data?

- 28.** You are performing a hypothesis test of a single population mean using a Student's t -distribution. The data are not from a simple random sample. Can you accurately perform the hypothesis test?
- 29.** You are performing a hypothesis test of a single population proportion. What must be true about the quantities of np and nq ?
- 30.** You are performing a hypothesis test of a single population proportion. You find out that np is less than five. What must you do to be able to perform a valid hypothesis test?
- 31.** You are performing a hypothesis test of a single population proportion. The data come from which distribution?

9.4 Rare Events, the Sample, Decision and Conclusion

- 32.** When do you reject the null hypothesis?
- 33.** The probability of winning the grand prize at a particular carnival game is 0.005. Is the outcome of winning very likely or very unlikely?
- 34.** The probability of winning the grand prize at a particular carnival game is 0.005. Michele wins the grand prize. Is this considered a rare or common event? Why?
- 35.** It is believed that the mean height of high school students who play basketball on the school team is 73 inches with a standard deviation of 1.8 inches. A random sample of 40 players is chosen. The sample mean was 71 inches, and the sample standard deviation was 1.5 years. Do the data support the claim that the mean height is less than 73 inches? The p -value is almost zero. State the null and alternative hypotheses and interpret the p -value.
- 36.** The mean age of graduate students at a University is at most 31 years with a standard deviation of two years. A random sample of 15 graduate students is taken. The sample mean is 32 years and the sample standard deviation is three years. Are the data significant at the 1% level? The p -value is 0.0264. State the null and alternative hypotheses and interpret the p -value.
- 37.** Does the shaded region represent a low or a high p -value compared to a level of significance of 1%?

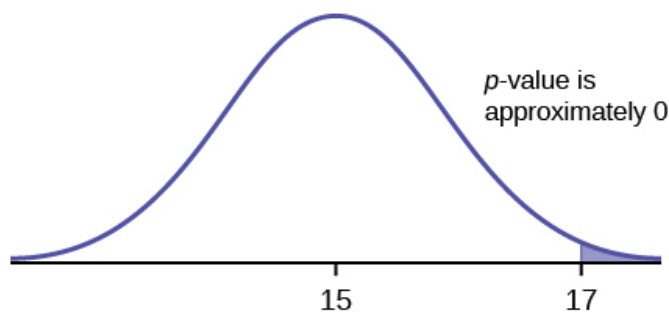


Figure 9.24

- 38.** What should you do when $\alpha > p$ -value?
- 39.** What should you do if $\alpha = p$ -value?
- 40.** If you do not reject the null hypothesis, then it must be true. Is this statement correct? State why or why not in complete sentences.

Use the following information to answer the next seven exercises: Suppose that a recent article stated that the mean time spent in jail by a first-time convicted burglar is 2.5 years. A study was then done to see if the mean time has increased in the new century. A random sample of 26 first-time convicted burglars in a recent year was picked. The mean length of time in jail from the survey was three years with a standard deviation of 1.8 years. Suppose that it is somehow known that the population standard deviation is 1.5. Conduct a hypothesis test to determine if the mean length of jail time has increased. Assume the distribution of the jail times is approximately normal.

- 41.** Is this a test of means or proportions?
- 42.** What symbol represents the random variable for this test?
- 43.** In words, define the random variable for this test.
- 44.** Is the population standard deviation known and, if so, what is it?
- 45.** Calculate the following:
- \bar{x} _____
 - σ _____

- c. s_x _____
- d. n _____

46. Since both σ and s_x are given, which should be used? In one to two complete sentences, explain why.
47. State the distribution to use for the hypothesis test.
48. A random survey of 75 death row inmates revealed that the mean length of time on death row is 17.4 years with a standard deviation of 6.3 years. Conduct a hypothesis test to determine if the population mean time on death row could likely be 15 years.
- a. Is this a test of one mean or proportion?
 - b. State the null and alternative hypotheses.
 H_0 : _____ H_a : _____
 - c. Is this a right-tailed, left-tailed, or two-tailed test?
 - d. What symbol represents the random variable for this test?
 - e. In words, define the random variable for this test.
 - f. Is the population standard deviation known and, if so, what is it?
 - g. Calculate the following:
 - i. \bar{x} = _____
 - ii. s = _____
 - iii. n = _____
 - h. Which test should be used?
 - i. State the distribution to use for the hypothesis test.
 - j. Find the p -value.
 - k. At a pre-conceived $\alpha = 0.05$, what is your:
 - i. Decision:
 - ii. Reason for the decision:
 - iii. Conclusion (write out in a complete sentence):

9.5 Additional Information and Full Hypothesis Test Examples

49. Assume $H_0: \mu = 9$ and $H_a: \mu < 9$. Is this a left-tailed, right-tailed, or two-tailed test?
50. Assume $H_0: \mu \leq 6$ and $H_a: \mu > 6$. Is this a left-tailed, right-tailed, or two-tailed test?
51. Assume $H_0: p = 0.25$ and $H_a: p \neq 0.25$. Is this a left-tailed, right-tailed, or two-tailed test?
52. Draw the general graph of a left-tailed test.
53. Draw the graph of a two-tailed test.
54. A bottle of water is labeled as containing 16 fluid ounces of water. You believe it is less than that. What type of test would you use?
55. Your friend claims that his mean golf score is 63. You want to show that it is higher than that. What type of test would you use?
56. A bathroom scale claims to be able to identify correctly any weight within a pound. You think that it cannot be that accurate. What type of test would you use?
57. You flip a coin and record whether it shows heads or tails. You know the probability of getting heads is 50%, but you think it is less for this particular coin. What type of test would you use?
58. If the alternative hypothesis has a not equals (\neq) symbol, you know to use which type of test?
59. Assume the null hypothesis states that the mean is at least 18. Is this a left-tailed, right-tailed, or two-tailed test?
60. Assume the null hypothesis states that the mean is at most 12. Is this a left-tailed, right-tailed, or two-tailed test?
61. Assume the null hypothesis states that the mean is equal to 88. The alternative hypothesis states that the mean is not equal to 88. Is this a left-tailed, right-tailed, or two-tailed test?

HOMEWORK

9.1 Null and Alternative Hypotheses

62. Some of the following statements refer to the null hypothesis, some to the alternate hypothesis.
 State the null hypothesis, H_0 , and the alternative hypothesis, H_a , in terms of the appropriate parameter (μ or p).

- a. The mean number of years Americans work before retiring is 34.
 - b. At most 60% of Americans vote in presidential elections.
 - c. The mean starting salary for San Jose State University graduates is at least \$100,000 per year.
 - d. Twenty-nine percent of high school seniors get drunk each month.
 - e. Fewer than 5% of adults ride the bus to work in Los Angeles.
 - f. The mean number of cars a person owns in her lifetime is not more than ten.
 - g. About half of Americans prefer to live away from cities, given the choice.
 - h. Europeans have a mean paid vacation each year of six weeks.
 - i. The chance of developing breast cancer is under 11% for women.
 - j. Private universities' mean tuition cost is more than \$20,000 per year.
- 63.** Over the past few decades, public health officials have examined the link between weight concerns and teen girls' smoking. Researchers surveyed a group of 273 randomly selected teen girls living in Massachusetts (between 12 and 15 years old). After four years the girls were surveyed again. Sixty-three said they smoked to stay thin. Is there good evidence that more than thirty percent of the teen girls smoke to stay thin? The alternative hypothesis is:
- a. $p < 0.30$
 - b. $p \leq 0.30$
 - c. $p \geq 0.30$
 - d. $p > 0.30$
- 64.** A statistics instructor believes that fewer than 20% of Evergreen Valley College (EVC) students attended the opening night midnight showing of the latest Harry Potter movie. She surveys 84 of her students and finds that 11 attended the midnight showing. An appropriate alternative hypothesis is:
- a. $p = 0.20$
 - b. $p > 0.20$
 - c. $p < 0.20$
 - d. $p \leq 0.20$
- 65.** Previously, an organization reported that teenagers spent 4.5 hours per week, on average, on the phone. The organization thinks that, currently, the mean is higher. Fifteen randomly chosen teenagers were asked how many hours per week they spend on the phone. The sample mean was 4.75 hours with a sample standard deviation of 2.0. Conduct a hypothesis test. The null and alternative hypotheses are:
- a. $H_0: \bar{x} = 4.5, H_a: \bar{x} > 4.5$
 - b. $H_0: \mu \geq 4.5, H_a: \mu < 4.5$
 - c. $H_0: \mu = 4.75, H_a: \mu > 4.75$
 - d. $H_0: \mu = 4.5, H_a: \mu > 4.5$

9.2 Outcomes and the Type I and Type II Errors

- 66.** State the Type I and Type II errors in complete sentences given the following statements.
- a. The mean number of years Americans work before retiring is 34.
 - b. At most 60% of Americans vote in presidential elections.
 - c. The mean starting salary for San Jose State University graduates is at least \$100,000 per year.
 - d. Twenty-nine percent of high school seniors get drunk each month.
 - e. Fewer than 5% of adults ride the bus to work in Los Angeles.
 - f. The mean number of cars a person owns in his or her lifetime is not more than ten.
 - g. About half of Americans prefer to live away from cities, given the choice.
 - h. Europeans have a mean paid vacation each year of six weeks.
 - i. The chance of developing breast cancer is under 11% for women.
 - j. Private universities mean tuition cost is more than \$20,000 per year.
- 67.** For statements a-j in **Exercise 9.109**, answer the following in complete sentences.
- a. State a consequence of committing a Type I error.
 - b. State a consequence of committing a Type II error.
- 68.** When a new drug is created, the pharmaceutical company must subject it to testing before receiving the necessary permission from the Food and Drug Administration (FDA) to market the drug. Suppose the null hypothesis is “the drug is unsafe.” What is the Type II Error?
- a. To conclude the drug is safe when in, fact, it is unsafe.
 - b. Not to conclude the drug is safe when, in fact, it is safe.
 - c. To conclude the drug is safe when, in fact, it is safe.
 - d. Not to conclude the drug is unsafe when, in fact, it is unsafe.

- 69.** A statistics instructor believes that fewer than 20% of Evergreen Valley College (EVC) students attended the opening midnight showing of the latest Harry Potter movie. She surveys 84 of her students and finds that 11 of them attended the midnight showing. The Type I error is to conclude that the percent of EVC students who attended is _____.
 a. at least 20%, when in fact, it is less than 20%.
 b. 20%, when in fact, it is 20%.
 c. less than 20%, when in fact, it is at least 20%.
 d. less than 20%, when in fact, it is less than 20%.

70. It is believed that Lake Tahoe Community College (LTCC) Intermediate Algebra students get less than seven hours of sleep per night, on average. A survey of 22 LTCC Intermediate Algebra students generated a mean of 7.24 hours with a standard deviation of 1.93 hours. At a level of significance of 5%, do LTCC Intermediate Algebra students get less than seven hours of sleep per night, on average?

The Type II error is not to reject that the mean number of hours of sleep LTCC students get per night is at least seven when, in fact, the mean number of hours

- a. is more than seven hours.
- b. is at most seven hours.
- c. is at least seven hours.
- d. is less than seven hours.

71. Previously, an organization reported that teenagers spent 4.5 hours per week, on average, on the phone. The organization thinks that, currently, the mean is higher. Fifteen randomly chosen teenagers were asked how many hours per week they spend on the phone. The sample mean was 4.75 hours with a sample standard deviation of 2.0. Conduct a hypothesis test, the Type I error is:

- a. to conclude that the current mean hours per week is higher than 4.5, when in fact, it is higher
- b. to conclude that the current mean hours per week is higher than 4.5, when in fact, it is the same
- c. to conclude that the mean hours per week currently is 4.5, when in fact, it is higher
- d. to conclude that the mean hours per week currently is no higher than 4.5, when in fact, it is not higher

9.3 Distribution Needed for Hypothesis Testing

72. It is believed that Lake Tahoe Community College (LTCC) Intermediate Algebra students get less than seven hours of sleep per night, on average. A survey of 22 LTCC Intermediate Algebra students generated a mean of 7.24 hours with a standard deviation of 1.93 hours. At a level of significance of 5%, do LTCC Intermediate Algebra students get less than seven hours of sleep per night, on average? The distribution to be used for this test is $\bar{X} \sim$ _____

- a. $N(7.24, \frac{1.93}{\sqrt{22}})$
- b. $N(7.24, 1.93)$
- c. t_{22}
- d. t_{21}

9.4 Rare Events, the Sample, Decision and Conclusion

73. The National Institute of Mental Health published an article stating that in any one-year period, approximately 9.5 percent of American adults suffer from depression or a depressive illness. Suppose that in a survey of 100 people in a certain town, seven of them suffered from depression or a depressive illness. Conduct a hypothesis test to determine if the true proportion of people in that town suffering from depression or a depressive illness is lower than the percent in the general adult American population.

- a. Is this a test of one mean or proportion?
- b. State the null and alternative hypotheses.
 H_0 : _____ H_a : _____
- c. Is this a right-tailed, left-tailed, or two-tailed test?
- d. What symbol represents the random variable for this test?
- e. In words, define the random variable for this test.
- f. Calculate the following:
 - i. $x =$ _____
 - ii. $n =$ _____
 - iii. $p' =$ _____
- g. Calculate $\sigma_x =$ _____. Show the formula set-up.
- h. State the distribution to use for the hypothesis test.
- i. Find the p -value.
- j. At a pre-conceived $\alpha = 0.05$, what is your:

- i. Decision:
- ii. Reason for the decision:
- iii. Conclusion (write out in a complete sentence):

9.5 Additional Information and Full Hypothesis Test Examples

For each of the word problems, use a solution sheet to do the hypothesis test. The solution sheet is found in **Appendix E**. Please feel free to make copies of the solution sheets. For the online version of the book, it is suggested that you copy the .doc or the .pdf files.

NOTE

If you are using a Student's- t distribution for one of the following homework problems, you may assume that the underlying population is normally distributed. (In general, you must first prove that assumption, however.)

74. A particular brand of tires claims that its deluxe tire averages at least 50,000 miles before it needs to be replaced. From past studies of this tire, the standard deviation is known to be 8,000. A survey of owners of that tire design is conducted. From the 28 tires surveyed, the mean lifespan was 46,500 miles with a standard deviation of 9,800 miles. Using $\alpha = 0.05$, is the data highly inconsistent with the claim?

75. From generation to generation, the mean age when smokers first start to smoke varies. However, the standard deviation of that age remains constant of around 2.1 years. A survey of 40 smokers of this generation was done to see if the mean starting age is at least 19. The sample mean was 18.1 with a sample standard deviation of 1.3. Do the data support the claim at the 5% level?

76. The cost of a daily newspaper varies from city to city. However, the variation among prices remains steady with a standard deviation of 20¢. A study was done to test the claim that the mean cost of a daily newspaper is \$1.00. Twelve costs yield a mean cost of 95¢ with a standard deviation of 18¢. Do the data support the claim at the 1% level?

77. An article in the *San Jose Mercury News* stated that students in the California state university system take 4.5 years, on average, to finish their undergraduate degrees. Suppose you believe that the mean time is longer. You conduct a survey of 49 students and obtain a sample mean of 5.1 with a sample standard deviation of 1.2. Do the data support your claim at the 1% level?

78. The mean number of sick days an employee takes per year is believed to be about ten. Members of a personnel department do not believe this figure. They randomly survey eight employees. The number of sick days they took for the past year are as follows: 12; 4; 15; 3; 11; 8; 6; 8. Let x = the number of sick days they took for the past year. Should the personnel team believe that the mean number is ten?

79. In 1955, *Life Magazine* reported that the 25 year-old mother of three worked, on average, an 80 hour week. Recently, many groups have been studying whether or not the women's movement has, in fact, resulted in an increase in the average work week for women (combining employment and at-home work). Suppose a study was done to determine if the mean work week has increased. 81 women were surveyed with the following results. The sample mean was 83; the sample standard deviation was ten. Does it appear that the mean work week has increased for women at the 5% level?

80. Your statistics instructor claims that 60 percent of the students who take her Elementary Statistics class go through life feeling more enriched. For some reason that she can't quite figure out, most people don't believe her. You decide to check this out on your own. You randomly survey 64 of her past Elementary Statistics students and find that 34 feel more enriched as a result of her class. Now, what do you think?

81. A Nissan Motor Corporation advertisement read, "The average man's I.Q. is 107. The average brown trout's I.Q. is 4. So why can't man catch brown trout?" Suppose you believe that the brown trout's mean I.Q. is greater than four. You catch 12 brown trout. A fish psychologist determines the I.Q.s as follows: 5; 4; 7; 3; 6; 4; 5; 3; 6; 3; 8; 5. Conduct a hypothesis test of your belief.

82. Refer to **Exercise 9.119**. Conduct a hypothesis test to see if your decision and conclusion would change if your belief were that the brown trout's mean I.Q. is **not** four.

83. According to an article in *Newsweek*, the natural ratio of girls to boys is 100:105. In China, the birth ratio is 100: 114 (46.7% girls). Suppose you don't believe the reported figures of the percent of girls born in China. You conduct a study. In this study, you count the number of girls and boys born in 150 randomly chosen recent births. There are 60 girls and 90 boys born of the 150. Based on your study, do you believe that the percent of girls born in China is 46.7%?

84. A poll done for *Newsweek* found that 13% of Americans have seen or sensed the presence of an angel. A contingent doubts that the percent is really that high. It conducts its own survey. Out of 76 Americans surveyed, only two had seen

or sensed the presence of an angel. As a result of the contingent's survey, would you agree with the *Newsweek* poll? In complete sentences, also give three reasons why the two polls might give different results.

85. The mean work week for engineers in a start-up company is believed to be about 60 hours. A newly hired engineer hopes that it's shorter. She asks ten engineering friends in start-ups for the lengths of their mean work weeks. Based on the results that follow, should she count on the mean work week to be shorter than 60 hours?

Data (length of mean work week): 70; 45; 55; 60; 65; 55; 55; 60; 50; 55.

86. Use the "Lap time" data for Lap 4 (see [Appendix C](#)) to test the claim that Terri finishes Lap 4, on average, in less than 129 seconds. Use all twenty races given.

87. Use the "Initial Public Offering" data (see [Appendix C](#)) to test the claim that the mean offer price was \$18 per share. Do not use all the data. Use your random number generator to randomly survey 15 prices.

NOTE

The following questions were written by past students. They are excellent problems!

88. "Asian Family Reunion," by Chau Nguyen

Every two years it comes around.

We all get together from different towns.

In my honest opinion,

It's not a typical family reunion.

Not forty, or fifty, or sixty,

But how about seventy companions!

The kids would play, scream, and shout

One minute they're happy, another they'll pout.

The teenagers would look, stare, and compare

From how they look to what they wear.

The men would chat about their business

That they make more, but never less.

Money is always their subject

And there's always talk of more new projects.

The women get tired from all of the chats

They head to the kitchen to set out the mats.

Some would sit and some would stand

Eating and talking with plates in their hands.

Then come the games and the songs

And suddenly, everyone gets along!

With all that laughter, it's sad to say

That it always ends in the same old way.

They hug and kiss and say "good-bye"

And then they all begin to cry!

I say that 60 percent shed their tears

But my mom counted 35 people this year.

She said that boys and men will always have their pride,

So we won't ever see them cry.

I myself don't think she's correct,

So could you please try this problem to see if you object?

89. "The Problem with Angels," by Cyndy Dowling

Although this problem is wholly mine,
 The catalyst came from the magazine, Time.
 On the magazine cover I did find
 The realm of angels tickling my mind.
 Inside, 69% I found to be
 In angels, Americans do believe.
 Then, it was time to rise to the task,
 Ninety-five high school and college students I did ask.
 Viewing all as one group,
 Random sampling to get the scoop.
 So, I asked each to be true,
 "Do you believe in angels?" Tell me, do!
 Hypothesizing at the start,
 Totally believing in my heart
 That the proportion who said yes
 Would be equal on this test.
 Lo and behold, seventy-three did arrive,
 Out of the sample of ninety-five.
 Now your job has just begun,
 Solve this problem and have some fun.

90. "Blowing Bubbles," by Sondra Prull

Studying stats just made me tense,
 I had to find some sane defense.
 Some light and lifting simple play
 To float my math anxiety away.
 Blowing bubbles lifts me high
 Takes my troubles to the sky.
 POIK! They're gone, with all my stress
 Bubble therapy is the best.
 The label said each time I blew
 The average number of bubbles would be at least 22.
 I blew and blew and this I found
 From 64 blows, they all are round!
 But the number of bubbles in 64 blows
 Varied widely, this I know.
 20 per blow became the mean
 They deviated by 6, and not 16.
 From counting bubbles, I sure did relax
 But now I give to you your task.
 Was 22 a reasonable guess?
 Find the answer and pass this test!

91. "Dalmatian Darnation," by Kathy Sparling

A greedy dog breeder named Spreckles

Bred puppies with numerous freckles
 The Dalmatians he sought
 Possessed spot upon spot
 The more spots, he thought, the more shekels.
 His competitors did not agree
 That freckles would increase the fee.
 They said, "Spots are quite nice
 But they don't affect price;
 One should breed for improved pedigree."
 The breeders decided to prove
 This strategy was a wrong move.
 Breeding only for spots
 Would wreak havoc, they thought.
 His theory they want to disprove.
 They proposed a contest to Spreckles
 Comparing dog prices to freckles.
 In records they looked up
 One hundred one pups:
 Dalmatians that fetched the most shekels.
 They asked Mr. Spreckles to name
 An average spot count he'd claim
 To bring in big bucks.
 Said Spreckles, "Well, shucks,
 It's for one hundred one that I aim."
 Said an amateur statistician
 Who wanted to help with this mission.
 "Twenty-one for the sample
 Standard deviation's ample:
 They examined one hundred and one
 Dalmatians that fetched a good sum.
 They counted each spot,
 Mark, freckle and dot
 And tallied up every one.
 Instead of one hundred one spots
 They averaged ninety six dots
 Can they muzzle Spreckles'
 Obsession with freckles
 Based on all the dog data they've got?

92. "Macaroni and Cheese, please!!!" by Nedda Misherghi and Rachelle Hall

As a poor starving student I don't have much money to spend for even the bare necessities. So my favorite and main staple food is macaroni and cheese. It's high in taste and low in cost and nutritional value.

One day, as I sat down to determine the meaning of life, I got a serious craving for this, oh, so important, food of my life. So I went down the street to Greatway to get a box of macaroni and cheese, but it was SO expensive! \$2.02 !!! Can you believe it? It made me stop and think. The world is changing fast. I had thought that the mean cost of a box (the normal size, not some super-gigantic-family-value-pack) was at most \$1, but now I wasn't so sure. However, I was determined to find out.

I went to 53 of the closest grocery stores and surveyed the prices of macaroni and cheese. Here are the data I wrote in my notebook:

Price per box of Mac and Cheese:

- 5 stores @ \$2.02
- 15 stores @ \$0.25
- 3 stores @ \$1.29
- 6 stores @ \$0.35
- 4 stores @ \$2.27
- 7 stores @ \$1.50
- 5 stores @ \$1.89
- 8 stores @ 0.75.

I could see that the cost varied but I had to sit down to figure out whether or not I was right. If it does turn out that this mouth-watering dish is at most \$1, then I'll throw a big cheesy party in our next statistics lab, with enough macaroni and cheese for just me. (After all, as a poor starving student I can't be expected to feed our class of animals!)

93. "William Shakespeare: The Tragedy of Hamlet, Prince of Denmark," by Jacqueline Ghodsi THE CHARACTERS (in order of appearance):

- HAMLET, Prince of Denmark and student of Statistics
- POLONIUS, Hamlet's tutor
- HORATIO, friend to Hamlet and fellow student

Scene: The great library of the castle, in which Hamlet does his lessons

Act I

(The day is fair, but the face of Hamlet is clouded. He paces the large room. His tutor, Polonius, is reprimanding Hamlet regarding the latter's recent experience. Horatio is seated at the large table at right stage.)

POLONIUS: My Lord, how can'st thou admit that thou hast seen a ghost! It is but a figment of your imagination!

HAMLET: I beg to differ; I know of a certainty that five-and-seventy in one hundred of us, condemned to the whips and scorns of time as we are, have gazed upon a spirit of health, or goblin damn'd, be their intents wicked or charitable.

POLONIUS If thou doest insist upon thy wretched vision then let me invest your time; be true to thy work and speak to me through the reason of the null and alternate hypotheses. (He turns to Horatio.) Did not Hamlet himself say, "What piece of work is man, how noble in reason, how infinite in faculties? Then let not this foolishness persist. Go, Horatio, make a survey of three-and-sixty and discover what the true proportion be. For my part, I will never succumb to this fantasy, but deem man to be devoid of all reason should thy proposal of at least five-and-seventy in one hundred hold true.

HORATIO (to Hamlet): What should we do, my Lord?

HAMLET: Go to thy purpose, Horatio.

HORATIO: To what end, my Lord?

HAMLET: That you must teach me. But let me conjure you by the rights of our fellowship, by the consonance of our youth, but the obligation of our ever-preserved love, be even and direct with me, whether I am right or no.

(Horatio exits, followed by Polonius, leaving Hamlet to ponder alone.)

Act II

(The next day, Hamlet awaits anxiously the presence of his friend, Horatio. Polonius enters and places some books upon the table just a moment before Horatio enters.)

POLONIUS: So, Horatio, what is it thou didst reveal through thy deliberations?

HORATIO: In a random survey, for which purpose thou thyself sent me forth, I did discover that one-and-forty believe fervently that the spirits of the dead walk with us. Before my God, I might not this believe, without the sensible and true avouch of mine own eyes.

POLONIUS: Give thine own thoughts no tongue, Horatio. (Polonius turns to Hamlet.) But look to't I charge you, my Lord. Come Horatio, let us go together, for this is not our test. (Horatio and Polonius leave together.)

HAMLET: To reject, or not reject, that is the question: whether 'tis nobler in the mind to suffer the slings and arrows of outrageous statistics, or to take arms against a sea of data, and, by opposing, end them. (Hamlet resignedly attends to his task.)

(Curtain falls)

94. "Untitled," by Stephen Chen

I've often wondered how software is released and sold to the public. Ironically, I work for a company that sells products with known problems. Unfortunately, most of the problems are difficult to create, which makes them difficult to fix. I usually

use the test program X, which tests the product, to try to create a specific problem. When the test program is run to make an error occur, the likelihood of generating an error is 1%.

So, armed with this knowledge, I wrote a new test program Y that will generate the same error that test program X creates, but more often. To find out if my test program is better than the original, so that I can convince the management that I'm right, I ran my test program to find out how often I can generate the same error. When I ran my test program 50 times, I generated the error twice. While this may not seem much better, I think that I can convince the management to use my test program instead of the original test program. Am I right?

95. "Japanese Girls' Names"

by Kumi Furuichi

It used to be very typical for Japanese girls' names to end with "ko." (The trend might have started around my grandmothers' generation and its peak might have been around my mother's generation.) "Ko" means "child" in Chinese characters. Parents would name their daughters with "ko" attaching to other Chinese characters which have meanings that they want their daughters to become, such as Sachiko—happy child, Yoshiko—a good child, Yasuko—a healthy child, and so on.

However, I noticed recently that only two out of nine of my Japanese girlfriends at this school have names which end with "ko." More and more, parents seem to have become creative, modernized, and, sometimes, westernized in naming their children.

I have a feeling that, while 70 percent or more of my mother's generation would have names with "ko" at the end, the proportion has dropped among my peers. I wrote down all my Japanese friends', ex-classmates', co-workers, and acquaintances' names that I could remember. Following are the names. (Some are repeats.) Test to see if the proportion has dropped for this generation.

Ai, Akemi, Akiko, Ayumi, Chiaki, Chie, Eiko, Eri, Eriko, Fumiko, Harumi, Hitomi, Hiroko, Hiroko, Hidemi, Hisako, Hinako, Izumi, Junko, Kana, Kanako, Kanayo, Kayo, Kayoko, Kazumi, Keiko, Keiko, Kei, Kumi, Kumiko, Kyoko, Kyoko, Madoka, Maho, Mai, Maiko, Maki, Miki, Miki, Mikiko, Mina, Minako, Miyako, Momoko, Nana, Naoko, Naoko, Naoko, Noriko, Rieko, Rika, Rika, Rumiko, Rei, Reiko, Reiko, Sachiko, Sachiko, Sachiyo, Saki, Sayaka, Sayoko, Sayuri, Seiko, Shiho, Shizuka, Sumiko, Takako, Takako, Tomoe, Tomoe, Tomoko, Touko, Yasuko, Yasuko, Yasuyo, Yoko, Yoko, Yoko, Yoshiko, Yoshiko, Yoshiko, Yuka, Yuki, Yuki, Yukiko, Yuko, Yuko.

96. "Phillip's Wish," by Suzanne Osorio

My nephew likes to play

Chasing the girls makes his day.

He asked his mother

If it is okay

To get his ear pierced.

She said, "No way!"

To poke a hole through your ear,

Is not what I want for you, dear.

He argued his point quite well,

Says even my macho pal, Mel,

Has gotten this done.

It's all just for fun.

C'mon please, mom, please, what the hell.

Again Phillip complained to his mother,

Saying half his friends (including their brothers)

Are piercing their ears

And they have no fears

He wants to be like the others.

She said, "I think it's much less.

We must do a hypothesis test.

And if you are right,

I won't put up a fight.

But, if not, then my case will rest.”

We proceeded to call fifty guys

To see whose prediction would fly.

Nineteen of the fifty

Said piercing was nifty

And earrings they'd occasionally buy.

Then there's the other thirty-one,

Who said they'd never have this done.

So now this poem's finished.

Will his hopes be diminished,

Or will my nephew have his fun?

97. "The Craven," by Mark Salangsang

Once upon a morning dreary

In stats class I was weak and weary.

Pondering over last night's homework

Whose answers were now on the board

This I did and nothing more.

While I nodded nearly napping

Suddenly, there came a tapping.

As someone gently rapping,

Rapping my head as I snore.

Quoth the teacher, "Sleep no more."

"In every class you fall asleep,"

The teacher said, his voice was deep.

"So a tally I've begun to keep

Of every class you nap and snore.

The percentage being forty-four."

"My dear teacher I must confess,

While sleeping is what I do best.

The percentage, I think, must be less,

A percentage less than forty-four."

This I said and nothing more.

"We'll see," he said and walked away,

And fifty classes from that day

He counted till the month of May

The classes in which I napped and snored.

The number he found was twenty-four.

At a significance level of 0.05,

Please tell me am I still alive?

Or did my grade just take a dive

Plunging down beneath the floor?

Upon thee I hereby implore.

98. Toastmasters International cites a report by Gallop Poll that 40% of Americans fear public speaking. A student believes that less than 40% of students at her school fear public speaking. She randomly surveys 361 schoolmates and finds that 135 report they fear public speaking. Conduct a hypothesis test to determine if the percent at her school is less than 40%.

99. Sixty-eight percent of online courses taught at community colleges nationwide were taught by full-time faculty. To test if 68% also represents California's percent for full-time faculty teaching the online classes, Long Beach City College (LBCC) in California, was randomly selected for comparison. In the same year, 34 of the 44 online courses LBCC offered were taught by full-time faculty. Conduct a hypothesis test to determine if 68% represents California. NOTE: For more accurate results, use more California community colleges and this past year's data.

100. According to an article in *Bloomberg Businessweek*, New York City's most recent adult smoking rate is 14%. Suppose that a survey is conducted to determine this year's rate. Nine out of 70 randomly chosen N.Y. City residents reply that they smoke. Conduct a hypothesis test to determine if the rate is still 14% or if it has decreased.

101. The mean age of De Anza College students in a previous term was 26.6 years old. An instructor thinks the mean age for online students is older than 26.6. She randomly surveys 56 online students and finds that the sample mean is 29.4 with a standard deviation of 2.1. Conduct a hypothesis test.

102. Registered nurses earned an average annual salary of \$69,110. For that same year, a survey was conducted of 41 California registered nurses to determine if the annual salary is higher than \$69,110 for California nurses. The sample average was \$71,121 with a sample standard deviation of \$7,489. Conduct a hypothesis test.

103. La Leche League International reports that the mean age of weaning a child from breastfeeding is age four to five worldwide. In America, most nursing mothers wean their children much earlier. Suppose a random survey is conducted of 21 U.S. mothers who recently weaned their children. The mean weaning age was nine months ($3/4$ year) with a standard deviation of 4 months. Conduct a hypothesis test to determine if the mean weaning age in the U.S. is less than four years old.

104. Over the past few decades, public health officials have examined the link between weight concerns and teen girls' smoking. Researchers surveyed a group of 273 randomly selected teen girls living in Massachusetts (between 12 and 15 years old). After four years the girls were surveyed again. Sixty-three said they smoked to stay thin. Is there good evidence that more than thirty percent of the teen girls smoke to stay thin?

After conducting the test, your decision and conclusion are

- Reject H_0 : There is sufficient evidence to conclude that more than 30% of teen girls smoke to stay thin.
- Do not reject H_0 : There is not sufficient evidence to conclude that less than 30% of teen girls smoke to stay thin.
- Do not reject H_0 : There is not sufficient evidence to conclude that more than 30% of teen girls smoke to stay thin.
- Reject H_0 : There is sufficient evidence to conclude that less than 30% of teen girls smoke to stay thin.

105. A statistics instructor believes that fewer than 20% of Evergreen Valley College (EVC) students attended the opening night midnight showing of the latest Harry Potter movie. She surveys 84 of her students and finds that 11 of them attended the midnight showing.

At a 1% level of significance, an appropriate conclusion is:

- There is insufficient evidence to conclude that the percent of EVC students who attended the midnight showing of Harry Potter is less than 20%.
- There is sufficient evidence to conclude that the percent of EVC students who attended the midnight showing of Harry Potter is more than 20%.
- There is sufficient evidence to conclude that the percent of EVC students who attended the midnight showing of Harry Potter is less than 20%.
- There is insufficient evidence to conclude that the percent of EVC students who attended the midnight showing of Harry Potter is at least 20%.

106. Previously, an organization reported that teenagers spent 4.5 hours per week, on average, on the phone. The organization thinks that, currently, the mean is higher. Fifteen randomly chosen teenagers were asked how many hours per week they spend on the phone. The sample mean was 4.75 hours with a sample standard deviation of 2.0. Conduct a hypothesis test.

At a significance level of $\alpha = 0.05$, what is the correct conclusion?

- There is enough evidence to conclude that the mean number of hours is more than 4.75
- There is enough evidence to conclude that the mean number of hours is more than 4.5
- There is not enough evidence to conclude that the mean number of hours is more than 4.5
- There is not enough evidence to conclude that the mean number of hours is more than 4.75

Instructions: For the following ten exercises,

Hypothesis testing: For the following ten exercises, answer each question.

- State the null and alternate hypothesis.
- State the p -value.
- State α .
- What is your decision?
- Write a conclusion.

f. Answer any other questions asked in the problem.

107. According to the Center for Disease Control website, in 2011 at least 18% of high school students have smoked a cigarette. An Introduction to Statistics class in Davies County, KY conducted a hypothesis test at the local high school (a medium sized—approximately 1,200 students—small city demographic) to determine if the local high school's percentage was lower. One hundred fifty students were chosen at random and surveyed. Of the 150 students surveyed, 82 have smoked. Use a significance level of 0.05 and using appropriate statistical evidence, conduct a hypothesis test and state the conclusions.

108. A recent survey in the *N.Y. Times Almanac* indicated that 48.8% of families own stock. A broker wanted to determine if this survey could be valid. He surveyed a random sample of 250 families and found that 142 owned some type of stock. At the 0.05 significance level, can the survey be considered to be accurate?

109. Driver error can be listed as the cause of approximately 54% of all fatal auto accidents, according to the American Automobile Association. Thirty randomly selected fatal accidents are examined, and it is determined that 14 were caused by driver error. Using $\alpha = 0.05$, is the AAA proportion accurate?

110. The US Department of Energy reported that 51.7% of homes were heated by natural gas. A random sample of 221 homes in Kentucky found that 115 were heated by natural gas. Does the evidence support the claim for Kentucky at the $\alpha = 0.05$ level in Kentucky? Are the results applicable across the country? Why?

111. For Americans using library services, the American Library Association claims that at most 67% of patrons borrow books. The library director in Owensboro, Kentucky feels this is not true, so she asked a local college statistic class to conduct a survey. The class randomly selected 100 patrons and found that 82 borrowed books. Did the class demonstrate that the percentage was higher in Owensboro, KY? Use $\alpha = 0.01$ level of significance. What is the possible proportion of patrons that do borrow books from the Owensboro Library?

112. The Weather Underground reported that the mean amount of summer rainfall for the northeastern US is at least 11.52 inches. Ten cities in the northeast are randomly selected and the mean rainfall amount is calculated to be 7.42 inches with a standard deviation of 1.3 inches. At the $\alpha = 0.05$ level, can it be concluded that the mean rainfall was below the reported average? What if $\alpha = 0.01$? Assume the amount of summer rainfall follows a normal distribution.

113. A survey in the *N.Y. Times Almanac* finds the mean commute time (one way) is 25.4 minutes for the 15 largest US cities. The Austin, TX chamber of commerce feels that Austin's commute time is less and wants to publicize this fact. The mean for 25 randomly selected commuters is 22.1 minutes with a standard deviation of 5.3 minutes. At the $\alpha = 0.10$ level, is the Austin, TX commute significantly less than the mean commute time for the 15 largest US cities?

114. A report by the Gallup Poll found that a woman visits her doctor, on average, at most 5.8 times each year. A random sample of 20 women results in these yearly visit totals

3; 2; 1; 3; 7; 2; 9; 4; 6; 6; 8; 0; 5; 6; 4; 2; 1; 3; 4; 1

At the $\alpha = 0.05$ level can it be concluded that the sample mean is higher than 5.8 visits per year?

115. According to the *N.Y. Times Almanac* the mean family size in the U.S. is 3.18. A sample of a college math class resulted in the following family sizes:

5; 4; 5; 4; 3; 6; 4; 3; 3; 5; 5; 6; 3; 3; 2; 7; 4; 5; 2; 2; 3; 2

At $\alpha = 0.05$ level, is the class' mean family size greater than the national average? Does the Almanac result remain valid? Why?

116. The student academic group on a college campus claims that freshman students study at least 2.5 hours per day, on average. One Introduction to Statistics class was skeptical. The class took a random sample of 30 freshman students and found a mean study time of 137 minutes with a standard deviation of 45 minutes. At $\alpha = 0.01$ level, is the student academic group's claim correct?

REFERENCES

9.1 Null and Alternative Hypotheses

Data from the National Institute of Mental Health. Available online at <http://www.nimh.nih.gov/publicat/depression.cfm>.

9.5 Additional Information and Full Hypothesis Test Examples

Data from Amit Schitai. Director of Instructional Technology and Distance Learning. LBCC.

Data from *Bloomberg Businessweek*. Available online at <http://www.businessweek.com/news/2011-09-15/nyc-smoking-rate-falls-to-record-low-of-14-bloomberg-says.html>.

Data from energy.gov. Available online at <http://energy.gov> (accessed June 27, 2013).

Data from Gallup®. Available online at www.gallup.com (accessed June 27, 2013).

Data from *Growing by Degrees* by Allen and Seaman.

Data from La Leche League International. Available online at <http://www.lalecheleague.org/Law/BAFeb01.html>.

Data from the American Automobile Association. Available online at www.aaa.com (accessed June 27, 2013).

Data from the American Library Association. Available online at www.ala.org (accessed June 27, 2013).

Data from the Bureau of Labor Statistics. Available online at <http://www.bls.gov/oes/current/oes291111.htm>.

Data from the Centers for Disease Control and Prevention. Available online at www.cdc.gov (accessed June 27, 2013).

Data from the U.S. Census Bureau, available online at <http://quickfacts.census.gov/qfd/states/00000.html> (accessed June 27, 2013).

Data from the United States Census Bureau. Available online at <http://www.census.gov/hhes/socdemo/language/>.

Data from Toastmasters International. Available online at <http://toastmasters.org/artisan/detail.asp?CategoryID=1&SubCategoryID=10&ArticleID=429&Page=1>.

Data from Weather Underground. Available online at www.wunderground.com (accessed June 27, 2013).

Federal Bureau of Investigations. “Uniform Crime Reports and Index of Crime in Daviess in the State of Kentucky enforced by Daviess County from 1985 to 2005.” Available online at <http://www.disastercenter.com/kentucky/crime/3868.htm> (accessed June 27, 2013).

“Foothill-De Anza Community College District.” De Anza College, Winter 2006. Available online at http://research.fhda.edu/factbook/DAdemofs/Fact_sheet_da_2006w.pdf.

Johansen, C., J. Boice, Jr., J. McLaughlin, J. Olsen. “Cellular Telephones and Cancer—a Nationwide Cohort Study in Denmark.” *Institute of Cancer Epidemiology and the Danish Cancer Society*, 93(3):203-7. Available online at <http://www.ncbi.nlm.nih.gov/pubmed/11158188> (accessed June 27, 2013).

Rape, Abuse & Incest National Network. “How often does sexual assault occur?” RAINN, 2009. Available online at <http://www.rainn.org/get-information/statistics/frequency-of-sexual-assault> (accessed June 27, 2013).

SOLUTIONS

1 The random variable is the mean Internet speed in Megabits per second.

3 The random variable is the mean number of children an American family has.

5 The random variable is the proportion of people picked at random in Times Square visiting the city.

7

a. $H_0: p = 0.42$

b. $H_a: p < 0.42$

9

a. $H_0: \mu = 15$

b. $H_a: \mu \neq 15$

11 Type I: The mean price of mid-sized cars is \$32,000, but we conclude that it is not \$32,000. Type II: The mean price of mid-sized cars is not \$32,000, but we conclude that it is \$32,000.

13 α = the probability that you think the bag cannot withstand -15 degrees F, when in fact it can β = the probability that you think the bag can withstand -15 degrees F, when in fact it cannot

15 Type I: The procedure will go well, but the doctors think it will not. Type II: The procedure will not go well, but the doctors think it will.

17 0.019

19 0.998

21 A normal distribution or a Student's t -distribution

23 Use a Student's t -distribution

25 a normal distribution for a single population mean

27 It must be approximately normally distributed.

29 They must both be greater than five.

31 binomial distribution

33 The outcome of winning is very unlikely.

35 $H_0: \mu \geq 73$

$H_a: \mu < 73$

The p -value is almost zero, which means there is sufficient data to conclude that the mean height of high school students who play basketball on the school team is less than 73 inches at the 5% level. The data do support the claim.

37 The shaded region shows a low p -value.

39 Do not reject H_0 .

41 means

43 the mean time spent in jail for 26 first time convicted burglars

45

- a. 3
- b. 1.5
- c. 1.8
- d. 26

47 $\bar{X} \sim N\left(2.5, \frac{1.5}{\sqrt{26}}\right)$

49 This is a left-tailed test.

51 This is a two-tailed test.

53

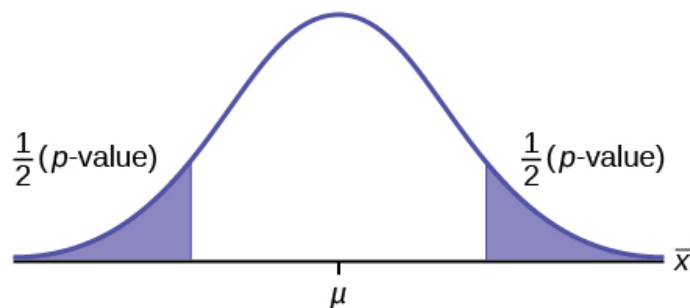


Figure 9.25

55 a right-tailed test

57 a left-tailed test

59 This is a left-tailed test.

61 This is a two-tailed test.

62

- a. $H_0: \mu = 34; H_a: \mu \neq 34$
- b. $H_0: p \leq 0.60; H_a: p > 0.60$
- c. $H_0: \mu \geq 100,000; H_a: \mu < 100,000$
- d. $H_0: p = 0.29; H_a: p \neq 0.29$
- e. $H_0: p = 0.05; H_a: p < 0.05$
- f. $H_0: \mu \leq 10; H_a: \mu > 10$
- g. $H_0: p = 0.50; H_a: p \neq 0.50$
- h. $H_0: \mu = 6; H_a: \mu \neq 6$
- i. $H_0: p \geq 0.11; H_a: p < 0.11$
- j. $H_0: \mu \leq 20,000; H_a: \mu > 20,000$

64 c**66**

- a. Type I error: We conclude that the mean is not 34 years, when it really is 34 years. Type II error: We conclude that the mean is 34 years, when in fact it really is not 34 years.
- b. Type I error: We conclude that more than 60% of Americans vote in presidential elections, when the actual percentage is at most 60%. Type II error: We conclude that at most 60% of Americans vote in presidential elections when, in fact, more than 60% do.
- c. Type I error: We conclude that the mean starting salary is less than \$100,000, when it really is at least \$100,000. Type II error: We conclude that the mean starting salary is at least \$100,000 when, in fact, it is less than \$100,000.
- d. Type I error: We conclude that the proportion of high school seniors who get drunk each month is not 29%, when it really is 29%. Type II error: We conclude that the proportion of high school seniors who get drunk each month is 29% when, in fact, it is not 29%.
- e. Type I error: We conclude that fewer than 5% of adults ride the bus to work in Los Angeles, when the percentage that do is really 5% or more. Type II error: We conclude that 5% or more adults ride the bus to work in Los Angeles when, in fact, fewer than 5% do.
- f. Type I error: We conclude that the mean number of cars a person owns in his or her lifetime is more than 10, when in reality it is not more than 10. Type II error: We conclude that the mean number of cars a person owns in his or her lifetime is not more than 10 when, in fact, it is more than 10.
- g. Type I error: We conclude that the proportion of Americans who prefer to live away from cities is not about half, though the actual proportion is about half. Type II error: We conclude that the proportion of Americans who prefer to live away from cities is half when, in fact, it is not half.
- h. Type I error: We conclude that the duration of paid vacations each year for Europeans is not six weeks, when in fact it is six weeks. Type II error: We conclude that the duration of paid vacations each year for Europeans is six weeks when, in fact, it is not.
- i. Type I error: We conclude that the proportion is less than 11%, when it is really at least 11%. Type II error: We conclude that the proportion of women who develop breast cancer is at least 11%, when in fact it is less than 11%.
- j. Type I error: We conclude that the average tuition cost at private universities is more than \$20,000, though in reality it is at most \$20,000. Type II error: We conclude that the average tuition cost at private universities is at most \$20,000 when, in fact, it is more than \$20,000.

68 b**70** d**72** d**74**

- a. $H_0: \mu \geq 50,000$

- b. $H_a: \mu < 50,000$
- c. Let \bar{X} = the average lifespan of a brand of tires.
- d. normal distribution
- e. $z = -2.315$
- f. $p\text{-value} = 0.0103$
- g. Check student's solution.
- h.
 - i. alpha: 0.05
 - ii. Decision: Reject the null hypothesis.
 - iii. Reason for decision: The p -value is less than 0.05.
 - iv. Conclusion: There is sufficient evidence to conclude that the mean lifespan of the tires is less than 50,000 miles.
- i. (43,537, 49,463)

76

- a. $H_0: \mu = \$1.00$
- b. $H_a: \mu \neq \$1.00$
- c. Let \bar{X} = the average cost of a daily newspaper.
- d. normal distribution
- e. $z = -0.866$
- f. $p\text{-value} = 0.3865$
- g. Check student's solution.
- h.
 - i. Alpha: 0.01
 - ii. Decision: Do not reject the null hypothesis.
 - iii. Reason for decision: The p -value is greater than 0.01.
 - iv. Conclusion: There is sufficient evidence to support the claim that the mean cost of daily papers is \$1. The mean cost could be \$1.
- i. (\$0.84, \$1.06)

78

- a. $H_0: \mu = 10$
- b. $H_a: \mu \neq 10$
- c. Let \bar{X} the mean number of sick days an employee takes per year.
- d. Student's t -distribution
- e. $t = -1.12$
- f. $p\text{-value} = 0.300$
- g. Check student's solution.
- h.
 - i. Alpha: 0.05
 - ii. Decision: Do not reject the null hypothesis.
 - iii. Reason for decision: The p -value is greater than 0.05.
 - iv. Conclusion: At the 5% significance level, there is insufficient evidence to conclude that the mean number of sick days is not ten.
- i. (4.9443, 11.806)

80

- a. $H_0: p \geq 0.6$
- b. $H_a: p < 0.6$
- c. Let P' = the proportion of students who feel more enriched as a result of taking Elementary Statistics.
- d. normal for a single proportion
- e. 1.12
- f. p -value = 0.1308
- g. Check student's solution.
- h.
 - i. Alpha: 0.05
 - ii. Decision: Do not reject the null hypothesis.
 - iii. Reason for decision: The p -value is greater than 0.05.
 - iv. Conclusion: There is insufficient evidence to conclude that less than 60 percent of her students feel more enriched.
- i. Confidence Interval: (0.409, 0.654)
The "plus-4s" confidence interval is (0.411, 0.648)

82

- a. $H_0: \mu = 4$
- b. $H_a: \mu \neq 4$
- c. Let \bar{X} the average I.Q. of a set of brown trout.
- d. two-tailed Student's t -test
- e. $t = 1.95$
- f. p -value = 0.076
- g. Check student's solution.
- h.
 - i. Alpha: 0.05
 - ii. Decision: Reject the null hypothesis.
 - iii. Reason for decision: The p -value is greater than 0.05
 - iv. Conclusion: There is insufficient evidence to conclude that the average IQ of brown trout is not four.
- i. (3.8865, 5.9468)

84

- a. $H_0: p \geq 0.13$
- b. $H_a: p < 0.13$
- c. Let P' = the proportion of Americans who have seen or sensed angels
- d. normal for a single proportion
- e. -2.688
- f. p -value = 0.0036
- g. Check student's solution.
- h.
 - i. alpha: 0.05
 - ii. Decision: Reject the null hypothesis.
 - iii. Reason for decision: The p -value is less than 0.05.
 - iv. Conclusion: There is sufficient evidence to conclude that the percentage of Americans who have seen or sensed an angel is less than 13%.
- i. (0, 0.0623).
The "plus-4s" confidence interval is (0.0022, 0.0978)

86

- a. $H_0: \mu \geq 129$
- b. $H_a: \mu < 129$
- c. Let \bar{X} = the average time in seconds that Terri finishes Lap 4.
- d. Student's t -distribution
- e. $t = 1.209$
- f. 0.8792
- g. Check student's solution.
- h.
 - i. Alpha: 0.05
 - ii. Decision: Do not reject the null hypothesis.
 - iii. Reason for decision: The p -value is greater than 0.05.
 - iv. Conclusion: There is insufficient evidence to conclude that Terri's mean lap time is less than 129 seconds.
- i. (128.63, 130.37)

88

- a. $H_0: p = 0.60$
- b. $H_a: p < 0.60$
- c. Let P' = the proportion of family members who shed tears at a reunion.
- d. normal for a single proportion
- e. -1.71
- f. 0.0438
- g. Check student's solution.
- h.
 - i. alpha: 0.05
 - ii. Decision: Reject the null hypothesis.
 - iii. Reason for decision: p -value < alpha
 - iv. Conclusion: At the 5% significance level, there is sufficient evidence to conclude that the proportion of family members who shed tears at a reunion is less than 0.60. However, the test is weak because the p -value and alpha are quite close, so other tests should be done.
- i. We are 95% confident that between 38.29% and 61.71% of family members will shed tears at a family reunion. (0.3829, 0.6171). The "plus-4s" confidence interval (see chapter 8) is (0.3861, 0.6139)

Note that here the "large-sample" 1 – PropZTest provides the approximate p -value of 0.0438. Whenever a p -value based on a normal approximation is close to the level of significance, the exact p -value based on binomial probabilities should be calculated whenever possible. This is beyond the scope of this course.

90

- a. $H_0: \mu \geq 22$
- b. $H_a: \mu < 22$
- c. Let \bar{X} = the mean number of bubbles per blow.
- d. Student's t -distribution
- e. -2.667
- f. p -value = 0.00486
- g. Check student's solution.
- h.
 - i. Alpha: 0.05
 - ii. Decision: Reject the null hypothesis.
 - iii. Reason for decision: The p -value is less than 0.05.

- iv. Conclusion: There is sufficient evidence to conclude that the mean number of bubbles per blow is less than 22.
- i. (18.501, 21.499)

92

- a. $H_0: \mu \leq 1$
- b. $H_a: \mu > 1$
- c. Let \bar{X} = the mean cost in dollars of macaroni and cheese in a certain town.
- d. Student's t -distribution
- e. $t = 0.340$
- f. $p\text{-value} = 0.36756$
- g. Check student's solution.
- h.
 - i. Alpha: 0.05
 - ii. Decision: Do not reject the null hypothesis.
 - iii. Reason for decision: The p -value is greater than 0.05
 - iv. Conclusion: The mean cost could be \$1, or less. At the 5% significance level, there is insufficient evidence to conclude that the mean price of a box of macaroni and cheese is more than \$1.
- i. (0.8291, 1.241)

94

- a. $H_0: p = 0.01$
- b. $H_a: p > 0.01$
- c. Let P' = the proportion of errors generated
- d. Normal for a single proportion
- e. 2.13
- f. 0.0165
- g. Check student's solution.
- h.
 - i. Alpha: 0.05
 - ii. Decision: Reject the null hypothesis
 - iii. Reason for decision: The p -value is less than 0.05.
 - iv. Conclusion: At the 5% significance level, there is sufficient evidence to conclude that the proportion of errors generated is more than 0.01.
- i. Confidence interval: (0, 0.094).
The "plus-4s" confidence interval is (0.004, 0.144).

96

- a. $H_0: p = 0.50$
- b. $H_a: p < 0.50$
- c. Let P' = the proportion of friends that has a pierced ear.
- d. normal for a single proportion
- e. -1.70
- f. $p\text{-value} = 0.0448$
- g. Check student's solution.
- h.
 - i. Alpha: 0.05
 - ii. Decision: Reject the null hypothesis
 - iii. Reason for decision: The p -value is less than 0.05. (However, they are very close.)

- iv. Conclusion: There is sufficient evidence to support the claim that less than 50% of his friends have pierced ears.
- i. Confidence Interval: (0.245, 0.515): The “plus-4s” confidence interval is (0.259, 0.519).

98

- a. $H_0: p = 0.40$
- b. $H_a: p < 0.40$
- c. Let P' = the proportion of schoolmates who fear public speaking.
- d. normal for a single proportion
- e. -1.01
- f. $p\text{-value} = 0.1563$
- g. Check student’s solution.
- h.
 - i. Alpha: 0.05
 - ii. Decision: Do not reject the null hypothesis.
 - iii. Reason for decision: The p -value is greater than 0.05.
 - iv. Conclusion: There is insufficient evidence to support the claim that less than 40% of students at the school fear public speaking.
- i. Confidence Interval: (0.3241, 0.4240): The “plus-4s” confidence interval is (0.3257, 0.4250).

100

- a. $H_0: p = 0.14$
- b. $H_a: p < 0.14$
- c. Let P' = the proportion of NYC residents that smoke.
- d. normal for a single proportion
- e. -0.2756
- f. $p\text{-value} = 0.3914$
- g. Check student’s solution.
- h.
 - i. alpha: 0.05
 - ii. Decision: Do not reject the null hypothesis.
 - iii. Reason for decision: The p -value is greater than 0.05.
 - iv. At the 5% significance level, there is insufficient evidence to conclude that the proportion of NYC residents who smoke is less than 0.14.
- i. Confidence Interval: (0.0502, 0.2070): The “plus-4s” confidence interval (see chapter 8) is (0.0676, 0.2297).

102

- a. $H_0: \mu = 69,110$
- b. $H_a: \mu > 69,110$
- c. Let \bar{X} = the mean salary in dollars for California registered nurses.
- d. Student's t -distribution
- e. $t = 1.719$
- f. $p\text{-value}: 0.0466$
- g. Check student’s solution.
- h.
 - i. Alpha: 0.05
 - ii. Decision: Reject the null hypothesis.
 - iii. Reason for decision: The p -value is less than 0.05.

- iv. Conclusion: At the 5% significance level, there is sufficient evidence to conclude that the mean salary of California registered nurses exceeds \$69,110.
- i. (\$68,757, \$73,485)

104 c

106 c

108

- a. $H_0: p = 0.488$ $H_a: p \neq 0.488$
- b. $p\text{-value} = 0.0114$
- c. $\alpha = 0.05$
- d. Reject the null hypothesis.
- e. At the 5% level of significance, there is enough evidence to conclude that 48.8% of families own stocks.
- f. The survey does not appear to be accurate.

110

- a. $H_0: p = 0.517$ $H_a: p \neq 0.517$
- b. $p\text{-value} = 0.9203$.
- c. $\alpha = 0.05$.
- d. Do not reject the null hypothesis.
- e. At the 5% significance level, there is not enough evidence to conclude that the proportion of homes in Kentucky that are heated by natural gas is 0.517.
- f. However, we cannot generalize this result to the entire nation. First, the sample's population is only the state of Kentucky. Second, it is reasonable to assume that homes in the extreme north and south will have extreme high usage and low usage, respectively. We would need to expand our sample base to include these possibilities if we wanted to generalize this claim to the entire nation.

112

- a. $H_0: \mu \geq 11.52$ $H_a: \mu < 11.52$
- b. $p\text{-value} = 0.000002$ which is almost 0.
- c. $\alpha = 0.05$.
- d. Reject the null hypothesis.
- e. At the 5% significance level, there is enough evidence to conclude that the mean amount of summer rain in the northeaster US is less than 11.52 inches, on average.
- f. We would make the same conclusion if α was 1% because the $p\text{-value}$ is almost 0.

114

- a. $H_0: \mu \leq 5.8$ $H_a: \mu > 5.8$
- b. $p\text{-value} = 0.9987$
- c. $\alpha = 0.05$
- d. Do not reject the null hypothesis.
- e. At the 5% level of significance, there is not enough evidence to conclude that a woman visits her doctor, on average, more than 5.8 times a year.

116

- a. $H_0: \mu \geq 150$ $H_a: \mu < 150$
- b. $p\text{-value} = 0.0622$
- c. $\alpha = 0.01$

- d. Do not reject the null hypothesis.
- e. At the 1% significance level, there is not enough evidence to conclude that freshmen students study less than 2.5 hours per day, on average.
- f. The student academic group's claim appears to be correct.